On Showing and Saying:

An Analysis of Gödel 1931

Submitted for Honors in Philosophy

Anthony MonCrief Coyne

To Thomas

Table of Contents

Preface

Mathematics is a constant source of valuable insight into philosophical problems. Many puzzles in philosophy possess analogues in mathematics for which acceptable solutions have been found. It is no surprise that Gödel's celebrated paper of 1931 is such a parallel. In his work, Gödel makes strong statements about a formal language from a formal proof expressed in that language. It is a situation comparable to that of the philosopher attempting to discuss language. Gödel's success in avoiding paradox points to what may be an important restriction on the capacities of language to talk.

When a specific example is used as a basis for conclusion, the writer is immediately open to the charge of hasty induction. This disucssion is not intended to demonstrate that the restriction developed is necessary. Indeed, if it is necessary, it may therefore be impossible to say so. Rather the force of the contrast between Richard's work and Gödel's is intended to provoke the realization that some restriction of the kind offered is needed.

The ideas developed in this paper have come from wide readings impossible to credit adequately. The most important of these sources appear in the Bibliography. Conversations with members of the Washington and Lee Philosophy Department have been most helpful. Mr. Robert Steck has been especially helpful in this regard. Occasional conversations with Mr. Gordon Williams have helped to assure me that I have not totally misunderstood Gödel's work. No research task of any magnitude is completed here without the assistance of either Mrs. Betty Munger or Miss Martha Cullipher. I thank them too.

Anthony MonCrief Coyne

Lexington, Virginia
April 27, 1970

On Showing and Saying:

An Analysis of Gödel 1931

1.1 In 1931 Kurt Gödel published a theorem demonstrating that many systems of formal logic capable of developing arithmetic possess major deficiencies. He shows that such systems, if they are consistent, are powerless to demonstrate their own consistency and possess propositions which can neither be proven nor disproven. In order for his results to obtain, Godel must use the language of his system to talk of itself. To avoid difficulties such as those which develop in Richard's paradox, Gödel must take care to insure his formal language refers only to numbers. The reader must then see this discussion as one about the system itself. From the consideration of this kind of language reference which keeps Gödel free from paradox, a general criterion for detecting paradoxical reference in language will be offered.

Traditional approaches to Gödel's work concern themselves with its relationship to the program of Hilbert for mathematics. Thus Jean Ladriere[1] in his study Limitations Internes des Formalismes examines the capacity of any formal system to embody mathematics in its entirety. The concern is mathematical. The concern of this paper is with language used for reference to objects (of some sort) beyond the language itself. Mathematical considerations, other than the proof itself, will be kept at a minimum.

One side issue has appeared in examining Gödel's Theorem deserving more attention than it has heretofore received. In his proof Gödel allows for expanding the scope of his system by allowing the inclusion of additional assumptions. In Chapter 3 the question of whether these postulates may be of empirical origin will be considered.

1.2 Chapter 2 of this paper is an exposition of the work of Kurt Gödel in his 1931 paper. It presents all of the technical devices used for the final result, omitting only proofs for some of the intermediate theorems which are of little interest in themselves. Chapter 3 develops Richard's paradox of 1905 along his original lines. It goes on to present the paradox as given by Nagel and Newman[2] and to develop the error which permits the paradox. While the two versions of the paradox are sufficiently dissimilar to permit two resolutions to the paradox, both explanations center around references by language which are beyond its permissable scope. Gödel's work is constrasted with Richard's to show that the difference between the two lies in Gödel's indirect reference. The structure of this indirect reference is considered as is the possibility for permitting empirical judgments as formulae. Finally Chaper 4 will explore a generalization of the solution of Richard's paradox for ordinary language. It will be shown that many traditional antinomies can be detected using this criterion.

II

An Exposition of Gödel's Theorem

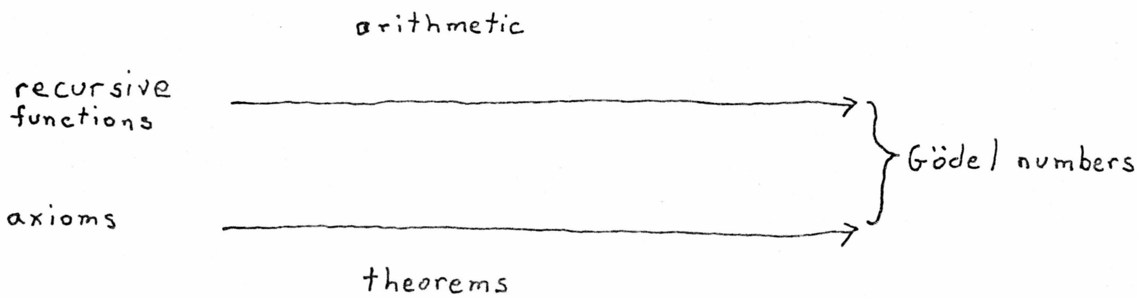
2.1 This section is a presentation of Godel's 1931 paper "Uber formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I."[1] While this section is not analytical in nature, the attention of the reader is called to the technique of Gödel numbering exhibited below and to the possible inclusion of an additional set of presumptions in the demonstration of Proposition VI. These will be the major subjects of the analytical chapter to follow.

There are two properties of a formal system of particular concern to the matatheory. These are consistency and completeness. A system is consistent if both a proposition and its negation cannot both be demonstrated. The system is called complete if every proposition can either be proven or disproven.

Gödel exhibits a well-formed formula of PM (Principia Mathematica) which says of itself that it is unprovable. This is shown, in addition, to be true by showing the assumption that either a proof or a disproof of the formula can be given leads to inconsistency. As a result PM will be shown to be incomplete.

There exist intelligible propositions which cannot be decided within the system itself. An immediate result of this is that if PM is consistent, it will be unable to demonstrate its own consistency.

Gödel proves his theorem by mirroring the system within its arithmetic. There are two levels in his demonstration, the level of the symbolic calculus, and the arithmetic which will represent it. The bridge to the arithmetic is made by assigning a number to each formula (Gödel numbering). To insure a precise copy of the symbolic system in its arithmetic, Gödel develops the notion of recursive functions so that he can talk about the numbers corresponding to the formulae in ways similar to the way he talks of the formulae. The development of the theorems follows from the axioms. The corresponding numerical transformation will be done with the recursive functions. The results Gödel obtains are demonstrated in the arithmetic about numbers. We then interpolate back to the theorems of the system to get his results.



(figure 1)

2.2 In carrying out his demonstration two techniques are employed. One is the use of Gödel numbering to provide indirect reference of the formulae to themselves. The other is the recursive function, a technical device which will be considered presently.

2.3 The recursive function is defined in two steps. The first step defines the function for an initial value. The second step defines the function at each point based on its value at the preceeding point. Consider the following recursive definition:

(1) (a) $f(1) = 5$
(b) $f(n+1) = f(n) + 3$

This set of formulae defines for every natural number a corresponding functional value. But it does so on the basis of a "recursive relation", (1b), which explains how f(n) is to be obtained from the value of f(n-1). The recursion begins with the setting of an initial value, here done by (1a). Thus f(1)=5, f(2)=f(1)+3, or 8. It is evident we may calculate the value of the function for the natural number of our choice, or check to see if f(1,234,687)=45,347,981,345,983. In practice considerable labor may be involved, but a schema for checking is available. The notion of recursive function is very similar to the notion of effective computability and to notions in the discussion of Turing machines. The function f above is said to be recursively defined by (1).

Gödel gives a more involved, but essentially similar, definition of recursive functions.[2] He states:

> A number theoretic function[3] $\phi(x_1, x_2, \dots x_n)$ is said to be recursively defined by the number theorectic functions $\psi(x_1, x_2, \dots x_{n-1})$ and $\mu(x_1, x_2, \dots x_{n+1})$, if for all $x_2, x_3, \dots x_n, k$ the following hold:
>
> $$\phi(0, x_2, \dots x_n) = \psi(x_2, \dots x_n)$$
> (2)
> $$\phi(k+1, x_2, \dots x_n) = \mu(k, \phi(k, x_2, \dots x_n), x_2, \dots x_n)$$
>
> A number theoretic function is called recursive, if there exist a finite series of number theoretic functions $\phi_1, \phi_2, \dots \phi_n$ which ends in ϕ and has the property that every function ϕ_k of the series is either recursively defined by two of the earlier ones, or is derived from any of the earlier ones by substitution, or, finally, is a constant or the successor function x+1.

If we look at Gödel's schema 2 we see it is our schema 1 expanded to accommodate several variables. If we suppress $x_2, \dots x_n$, we get

$$\phi(0) = \text{some constant}$$
$$\phi(k+1) = \mu(k, \phi(k))$$

which is our schema 1. In effect what has been done is to recursively define one of the variables in ϕ . The function ϕ is then said to be recursively defined if ψ and μ are recurisvely defined in terms of recursively defined functions, finitely back to the two primitive functions[4]

$g(n) = a$ where a is a constant

$f(n) = n+1$ the successor function

If ψ and μ admit of recursive definition, then, in theory at least, ϕ could be recursively defined solely in terms of the constant function and the successor function.

The notation is simpler if we allow these intermediate functions. It is possible to define addition recursively. Informally this means we possess techniques for checking addition by recourse to the notion of sequence. Further, only finitely many references to the sequence will be needed to carry out each check.

Similarly multiplication can be defined recursively and eventually many other functions. Each function so definied can have any particular of its values determined after a finite number of simple operations. All questions about the value of such functions thus has a schema for decision.

2.4 The other tool used by Gödel is one for indirect reference to the statements of the calculus. He talks about numbers but allows us to see this as talk about the syntax of the system by letting each statement (and each series of such statements) be represented by a natural number. In a way the entire set of statements is contained in the natural numbers of the system because of the correspondence between them which is established by the numbering scheme. Though it is not this simple, the following indicates the spirit of the attempt. Suppose every statement is represented by a natural number:

(1) All swans are white.
(2) All white things are pure.
(3) All swans are pure.

(1) and (2) together imply (3). Here it is also true that 1+2=3. In the arithmetic, addition represents the operation of implication at the symbolic level in the arithmetic level. If all statements could be numbered in such a way that if 5+135=140, then statement 5 and statement 135 imply statement 140, then we could dismiss the operation of implication between statements and discover all such relations by looking at addition. Logic would be a matter of arithmetic.

It is with this end in view that Gödel establishes his correspondence between numbers and statements. His schema is based on the fact that each non-prime number can be expressed as a product of primes in exactly one way, disregarding the order of the factors.[5] To the basic signs of his system Gödel assigns odd integers according to the following table:

Sign	Meaning	Number
0	zero	1
f	successor function	3
$\sim$	negation	5
$\vee$	disjunction	7
Π	universal quantification	9
(	left parenthesis	11
)	right parenthesis	13

For variables of type n the number is p^n where p is a prime over 13. Thus there are denumerably many variables of each type.

(figure 2)

We now use these numbers of signs as a basis on which to assign numbers to statements. We do so in the following way: to the sequence of numbers of signs $n_1, n_2, \ldots n_K$ we assign the number $2^{n_1} \times 3^{n_2} \times 5^{n_3} \ldots p_K^{n_K}$, where the p_i are increasing primes beginning with 2. For example, to the sequence of signs

$$\sim(a \vee b)$$

we assign the corresponding sequence of numbers

$$5, 11, 17, 7, 19, 13$$

Since a and b are assumed to be variables of the first type, they may be paired with 17 and 19, the first powers of primes greater than 13. Now to the expression as a whole we assign the number

$$2^5 \times 3^{11} \times 5^{17} \times 7^7 \times 11^{19} \times 13^{13}$$

Numbers can also be worked back to sequences of signs. Many numbers, however, do not represent such sequences. For example, $100 = 2^2 \times 5^2$ is not one since 2 and 5, the primes present as factors, are not successive primes and hence 100 could not have resulted from numbering according to our system. But

$$2^5 \times 3^{17}$$

represents $\sim x$, where x is the first variable of the lowest type. We can thus go from expressions to numbers and where it is the case, from numbers to corresponding expressions.[6]

2.5 The system which Gödel uses for his proof is a modified version of PM. Peano's axioms[7] have been added (though this only simplifies matters, since Peano's axioms can be deduced within PM). Specifically the axioms of Gödel's system are:

I. Peano's axioms[8]

1. zero is not the successor of any natural number

$$\sim (f(x_1) = 0)$$

2. two numbers which have the same successor are equal

$$f(x_1) = f(y_1) \supset x_1 = y_1$$

3. the principle of mathematical induction: if a property holds for zero and, if it holds for any natural number it holds for its successor, then the property holds for all natural numbers.

$$x_2(0) \cdot x_1\Pi(x_2(x_1) \supset x_2(f(x_1))) \supset x_1\Pi(x_2(x_1))$$

II. The following axiom forms (in which any formula may be substituted for p, q, or r to produce an axiom):

1. $p \vee p \supset p$
2. $p \supset p \vee q$
3. $p \vee q \supset q \vee p$
4. $(p \supset q) \supset (r \vee p \supset r \vee q)$

III. The following schema, which concern the quantificational calculus[9]:

1. $v\Pi(a) \supset \mathrm{Subst}\left(a \begin{smallmatrix} v \\ c \end{smallmatrix}\right)$, where c is a particular
2. $v\Pi(b \vee a) \supset b \vee v\Pi(a)$, if v does not occur free in b

IV. The axiom of reducibility is accepted in this form:

$$(Eu)(v\Pi(u(v) \equiv a))$$

V. The following axiom (together with all of its type-lifts) which state that a class is completely determined by its elements:

$$x_1\Pi(x_2(x_1) = y_2(x_1)) \supset x_2 = y_2$$

Fortunately these axioms will not be referred to again. They are presented to show on what a small base a system can be built and still be adequate to demonstrate Gödel's result.

With one more definition we can approach the proof itself. A formula c is called an immediate consequence of a and b, if a is the formula $(\sim(b)) \vee c$; c is an immediate consequence of a if c is the formula $v\Pi(a)$ where v is a variable of a. The provable formulae are the smallest class which contains all of the axioms and their consequences.

2.6 We are now ready to look at the proof proper. It begins with an exposition of recursive functions. These functions will turn out to be models of the concepts such as proof, entailment, etc. They thus provide formulae for dealing with the numbers which refer to the statements of the system.

The line of development is this. First, there will be five theorems about combinations of recursive functions. This will serve to justify the exposition of 45 recursive functions ending in

$$xBy \equiv Bw(x) \text{ and } [l(x)]\,Gl(x) = y$$

that is, x is the number corresponding to the proof of the formula whose number is y. This function is shown to be recursive. Then we are in position to prove Proposition VI which asserts the incompleteness of PM. Proposition XI on the inability of PM to demonstrate its consistency will then follow immediately.

The first five propositions about recursive functions will be stated without proof. Gödel sketches these in his paper. The first three will seem obvious, if the definition of recursive function is kept in mind. These five theorems tell us how we can combine recursive functions to produce recursive functions. They form a calculus for these functions.

2.7 I. Every function (or relation)[10] derived from recursive functions (or relations) by the substitution of recursive functions in place of variables is recursive; so also is every function derived from recursive functions by by recursive definition according to schema (2).

II. If R and S are recursive relations, then so are $\sim R$ (or $\bar{R}$), $R \vee S$ (and therefore $R \cdot S$).

III. If the functions $\phi(\mathfrak{x})$[11] and $\psi(\mathfrak{n})$ are recursive, so also is the relation $\phi(\mathfrak{x}) = \psi(\mathfrak{n})$.

IV. If the function $\phi(\mathfrak{x})$ and the relation $R(x, \mathfrak{n})$ are recursive, so also are the relations S,T

$$S(\mathfrak{x}, \mathfrak{n}) = (Ex)[x \leq \phi(\mathfrak{x}) \text{ and } R(x, \mathfrak{n})]$$

$$T(\mathfrak{x}, \mathfrak{n}) = (x)[x \leq \phi(\mathfrak{x}) \supset R(x, \mathfrak{n})]$$

and likewise the function

$$\psi(\mathfrak{x}, \mathfrak{n}) = \varepsilon x (x \leq \phi(\mathfrak{x}) \text{ and } R(x, \mathfrak{n}))$$

where ε is a function satisfied by the smallest integer x for which the following condition holds, and 0 if there is no such number.

This theorem will be needed in establishing the recursiveness of functions which express manipulations on particular terms in number sequences.

One more theorem remains to be shown about recursive functions. This one guarantees that a recursive function exists for each recursive relation. The conclusion of the theorem is given in the form of Gödel numbers. For a proof, see Gödel's paper.

Proposition V: To every recursive relation $R(x_1, \dots x_n)$ (a statement) there corresponds an n-place relation sign r (a number) (with free variables $v_1, v_2, \dots v_n$), such that for every n-tuple of numbers $(x_1, x_2, \dots x_n)$ the following hold:

$$R(x_1, \dots x_n) \rightarrow Bew[Sb(r \begin{smallmatrix} v_1 & \dots & v_n \\ Z(x_1) & \dots & Z(x_n) \end{smallmatrix})]$$
$$\bar{R}(x_1, \dots x_n) \rightarrow Bew[Neg\ Sb(r \begin{smallmatrix} v_1 & \dots & v_n \\ Z(x_1) & \dots & Z(x_n) \end{smallmatrix})]$$

The claim is that if a relation holds among n-variables, the number which corresponds to the relation when the variables are replaced is among the class of provable formulae. Similarly, the second proposition states the result when the relation does not hold among the numbers.

2.8 It can easily be shown that $x+y$, $x \cdot y$, x^y , $x < y$, $x = y$, are recursive. In the original paper 45 functions follow which are immediately seen to be recursive on the basis of the preceding theorems. They are functions on natural numbers which can be taken as referring to the signs of the system and to the sequence of such signs. The goal of these propositions is to establish that the notion of being a proof can be recursively defined. The result (number 45) is (B comes from the German Beweis):

$$x B y \equiv Bw(x) \text{ and } [l(x)]\, Gl\, x = y$$

Interpreted, x is the Gödel number of a series of formulae which constitute a proof schema, $Bw(x)$, and whose last term $[l(x)]\, Gl\, x$ is the Gödel number of statement y. Gödel next defines one more function which is not recursive, and which asserts that the number x corresponds to a provable formula. The function is number 46 and draws its name from the German Beweisbar:

$$Bew(x) = (Ey)\, y B x$$

Five more of these functions will be used in the final demonstration. These are Sb, Z, Gen, Neg, and Flg.

Sb is number 31 of the series and corresponds to the operation of substitution. It generates the Godel number which corresponds to the substitution of a constant in a variable of the function. Sb operates on numbers, performing the corresponding operation to variable substitution.

$$Sb\left(x \begin{smallmatrix} v \\ y \end{smallmatrix}\right)$$

$$\left.\begin{array}{l} x_v \xrightarrow{\quad Sb\left(x \begin{smallmatrix} v \\ y \end{smallmatrix}\right) \quad} x_y \\ \\ x(v) \xrightarrow{\quad \text{substitution} \quad} x(y) \end{array}\right\} \begin{array}{l} \text{arithmetic} \\ \text{Gödel numbers} \\ \text{statements} \end{array}$$

Z (from the German Zahlzeichen) provides for each number its Gödel number. Since numbers are not primitive in the system, they are defined in terms of the successor function f by expressions such as $f(f(o))$. Hence each number possesses a corresponding Gödel number.

The function Gen (number 15) corresponds to the logical operation of generalization. It, like Sb, operates on numbers, and corresponds to the manipulation of symbols in the system.

$$\left.\begin{array}{l} n_1 \xrightarrow{x\ Gen\ n_1} n_2 \\ \\ f(x) \xrightarrow[\text{Generalization}]{} (x)\, f(x) \end{array}\right\} \text{Gödel numbers}$$

Neg(x) is the negation. Flg(x), from the German Folgerungsmenge, is defined as a set of numbers-- the set which corresponds to the set of provable formulae.

<u>2.9</u> To be honest, the demonstration of Proposition VI will require a stronger demand be placed on the system than simple consistency. Simple consistency requires that p and $\sim p$ not be among the consequences of the axioms. ω -consistency, which Gödel will require[12] prohibits the following in the system:

1. Proposition x holds for each natural number.

2. It is not the case that proposition x holds for every natural number.

One is a claim about the numbers demonstrated individually.

Two states that the negation of the generalization is demonstrable.

Every ω -consistent system is simply consistent. However ω -consistency cannot be demonstrated from the assumption of simple consistency.[13] Formally Gödel gives this condition of ω -consistency as a statement about numbers. A system is ω-consistent if there is no class sign number a in c such that

$$(n)\left[Sb\left(a_{Z(n)}^{v}\right) \in Flg(c)\right] \text{ and } \left[Neg\,(v\,Gen\,a)\right] \in Flg(c)$$

2.10 And now as Gödel says, "we come to the object of our exercises":

Proposition VI: To every ω -consistent recursive class c of formulae there correspond recursive class-signs r such that neither $v\,Gen\,r$ nor $Neg(v\,Gen\,r)$ belongs to $Flg(c)$.

v is the free variable of r.

c is a set of formulae in addition to those given by the axioms. We will need ω -consistency and recursiveness for this class, but nothing else. It may be null (as it would be in applying Gödel's work to PM). The introduction of this additional set of propositions at this point will be analyzed in the discussion to follow.

For the new set we define $Bw_c(x)$ similar to the function Bw exhibited above. The difference between the two is in the acceptance of propositions from c as well as from the system itself.

$$Bw_c(x) = (n)\,[n \leq l(x) \rightarrow Ax(n\,Gl\,x) \vee (n\,Gl\,x)\,Gc \vee (Ep,q)\{0 < p,q < n \text{ and } Fl(n\,Gl\,x,\ p\,Gl\,x,\ q\,Gl\,x)\} \text{ and } l(x) > 0]$$

cf. with Gödel's function 44

Analogous to 45 and 46 we define:

$$x\,B_c\,y \equiv Bw_c(x) \text{ and } [l(x)]\,Gl\,x = y$$

$$Bew_c\,x \equiv (Ey)\,y\,B_c\,x$$

The following propositions are now evident:

(1) $(x)\,[Bew_c(x) \equiv x \in Flg(c)]$

(2) $(x)\,[Bew(x) \rightarrow Bew_c(x)]$

(1) states that if x can be proven on the basis of the expanded system, then it is among the consequences of the expanded system. (2) states that if a number x corresponds to a provable statement in the basic system, it also corresponds to one in the expanded system.

We now define the following recursive relation:

(3) $Q(x,y) = \overline{x\,B_c\,(Sb\,y^{19}_{Z(y)})}$

This relation will be the crucial function for our demonstration. Let us review its contents:

x, y are integers.

Q is a recursive relation defined on pairs of integers.

$x\, B_c\, y$ says that the series of statements corresponding to x are a proof for y.

y is a number corresponding to a statement of the calculus. It possibly contains free variables, and if it does, they would have numbers 17,19,23,27.... Thus we may substitute in variable 19 a number. We will substitute Z(y), the Godel number corresponding to the number y.

The bar sign is negation.

$$\overline{x\, B_c\, [Sb(y\, {}^{19}_{Z(y)})]}$$

The relation expresses the condition that x does not correspond to a proof for y, when y has one of its variables replaced by the Gödel number of its Gödel number.

Since Q is a recursive relation, there exists a function Q' which holds if and only if the relation Q holds. By proposition V the following holds about the Gödel number q corresponding to the function Q' which determines the relation Q:

$$(3) \quad \overline{x\, B_c\, [Sb(y\, {}^{19}_{Z(y)})]} \longrightarrow Bew_c\, [Sb(q\, {}^{17}_{Z(x)}\, {}^{19}_{Z(y)})]$$

$$(4) \quad x\, B_c\, [Sb(y\, {}^{19}_{Z(y)})] \longrightarrow Bew_c\, [Neg\, Sb(q\, {}^{17}_{Z(x)}\, {}^{19}_{Z(y)})]$$

If the relation Q holds between x and y, the function which generates Q holds when x and y are substituted, and thus the number q corresponding to the function after the substitution is a provable number. The second relation holds if $\sim Q(x,y)$. Note the negation, Q holds if x is not a proof.

Now consider the number

$$p = 17 \; Gen \; q$$

We substitute this number for the variable y in (3):

$$(5) \quad x \overline{B_c} \left[Sb \left(p \begin{smallmatrix} 19 \\ Z(p) \end{smallmatrix} \right) \right] \rightarrow Bew_c \left[Sb \left(q \begin{smallmatrix} 17 \\ Z(x) \end{smallmatrix} \begin{smallmatrix} 19 \\ Z(p) \end{smallmatrix} \right) \right]$$

In effect we have narrowed our consideration from all the formulae y to the formula represented by p. Expanding the left side of (5) we get:

$$Sb \left(p \begin{smallmatrix} 19 \\ Z(p) \end{smallmatrix} \right)$$

$$Sb \left(17 \; Gen \; q \begin{smallmatrix} 19 \\ Z(p) \end{smallmatrix} \right)$$

$$17 \; Gen \left(Sb \; q \begin{smallmatrix} 19 \\ Z(p) \end{smallmatrix} \right)$$

Sub and Gen are commutative when they refer to different variables. We call $Sb \, q \begin{smallmatrix} 19 \\ Z(p) \end{smallmatrix}$, r and thus the left side becomes:

$$x \overline{B_c} \; 17 \; Gen \; r$$

Shifting our attention to the right side we can change the form of the argument of Bew_c thus:

$$Sb\left(q \begin{smallmatrix}17 \\ Z(x)\end{smallmatrix} \begin{smallmatrix}19 \\ Z(p)\end{smallmatrix}\right)$$

$$Sb\left(q \begin{smallmatrix}19 \\ Z(p)\end{smallmatrix} \begin{smallmatrix}17 \\ Z(x)\end{smallmatrix}\right)$$

since the order of the substitution is immaterial, if the variables are different. Using the r defined above, this becomes:

$$Sb\left(r \begin{smallmatrix}17 \\ Z(x)\end{smallmatrix}\right)$$

Equations (3) and (4) can then be written as

$$(6) \quad x \overline{B}_c (17\, Gen\, r) \rightarrow Bew_c \left[Sb\left(r \begin{smallmatrix}17 \\ Z(x)\end{smallmatrix}\right)\right]$$

$$(7) \quad x B_c (17\, Gen\, r) \longrightarrow Bew_c \left[Neg\ Sb\left(r \begin{smallmatrix}17 \\ Z(x)\end{smallmatrix}\right)\right]$$

We immediately sense something amiss about this pair. (6) indicates that if x does not represent a proof for r holding for all x, then r holds for this x. (7) indicates that if x is a proof for r holding for all x, then x itself does not satisfy r.

We now inquire about the provability of $17\, Gen\, r$.

A. Suppose $17\, Gen\, r$ were provable. Then there would be a number x which represents its proof, that is $x B_c (17\, Gen\, r)$ would hold and thus from (7) $Bew_c\left[Neg\ Sb\left(r \begin{smallmatrix}17 \\ Z(x)\end{smallmatrix}\right)\right]$. Thus $Neg\ Sb\left(r \begin{smallmatrix}17 \\ Z(x)\end{smallmatrix}\right)$ holds since it can be proven. But $17\, Gen\, r$ holds by assumption, whence $Sb(r^{17}_{Z(x)})$ holds. Thus the system is inconsistent and therefore ω -inconsistent. Thus $17\ Gen\ r$ cannot be proven.

B. Suppose that $Neg\ (17\ Gen\ r)$ were provable. Then it must be true

$$(8) \quad (x)\ x\ \overline{B_c}\ (17\ Gen\ r)$$

for no x could represent a proof for $17\ Gen\ r$ and the system remain consistent. We generalize (6) as a statement for all x and get

$$(6') \quad (x)\ x\ \overline{B_c}\ (17\ Gen\ r) \rightarrow (x)\ Bew_c\left[Sb(r^{17}_{Z(x)})\right]$$

Thus from (8) and (6') we get

$$(9) \quad (x)\ Bew_c\left[Sb(r^{17}_{Z(x)})\right]$$

Now consider $Neg(17\ Gen\ r)$ together with (9). (9) asserts r holds for each x but $Neg\ 17\ Gen\ r$ asserts that it cannot hold for all x. Thus the system is ω -inconsistent if $Neg\ (17\ Gen\ r)$ can be proven.

The assumption that either $17\ Gen\ r$ or $Neg\ (17\ Gen\ r)$ is provable produces ω -inconsistency. Thus if the system is ω -consistent, $17\ Gen\ r$ must be undecidable.

2.11 Gödel's proposition on the non-provability of consistency follows immediately from the completeness result.

Proposition XI: If c be a given recursive, consistent class of formulae, then the propositional formula which states that c is consistent is not c-provable.

Since the unprovability of $17\,Gen\,r$ depended only on the consistency of the system (and not on the stronger ω -consistency) we have (where Wid is consistency):

$$Wid(c) \rightarrow \overline{Bew_c}\,(17\,Gen\,r)$$

since if the system were inconsistent, $17\,Gen\,r$ would be provable.

Thus:

$$Wid(c) \rightarrow (x)\, x\, \overline{B_c}\,(17\,Gen\,r)$$

Since $17\,Gen\,r = p$ and the function implied is the relation Q of Proposition VI, we may rewrite this as

(10) $\qquad Wid(c) \rightarrow (x)\; Q(x, p)$

Now all our techniques of proof have been arithmetic and can be expressed within the formal system. In particular $Wid(c)$ is expressible and has Gödel number w. $Q(x,y)$ is expressible in the system by q, and $Q(x,p)$ has number r, since $r = Sb\left(q\,{}^{19}_{Z(p)}\right)$.

Thus $(x)\,Q(x,p)$ has the number $17\,Gen\,r$. Thus $w\,Imp\,(17\,Gen\,r)$ corresponding to (10) is provable. Now w cannot be the Gödel number of a theorem, or $17\,Gen\,r$, the number of an undecidable proposition, would be the number of a provable formula. Therefore PM cannot establish its own consistency.

2.12 Proposition XI, together with Proposition VI constitute the heart of Gödel's 1931 paper. Together they place limits on deductive systems. VI indicates that many systems will always be incomplete. XI goes on to show that these same systems will be inadequate to demonstrate their own consistency. This last result questions the possibility of a consistency proof at all, since such a proof could never be imaged within a formal system which contains ordinary mathematics. Presumably such a proof would be based on a larger set of axioms and thus make the consistency of the system which demonstrated the consistency a pertinent question.

On Richard and Gödel: The Difference

3.1 In this chapter we shall explore by contrast Gödel's Theorem with Richard's Paradox. We shall see the only difference between the two ~~t~~o be Richard's insistence on speaking directly to the statements at issue as opposed to Gödel's technique of never making such reference. From this point, the relationship among observer, language, and object of language necessary for the success of the theorem will be exposed. In this context the observer will be seen to have a crucial role. Finally the theories of empirical science know by the observer alongside the system will be considered and it will be shown that this can have no effect on the apprehension the observer has of the theorem.

3.2 Richard's Paradox, first presented in 1905, concerns some difficulties in general set theory. It will be clear that Richard's technique is very close to Gödel's. In Richard's paradox a contradiction develops from a direct consideration by a system of itself.

We may assign a number to any finitely long English statement in a one to one fashion; for example, this may be done by first ordering the sentences by length and then in alphabetical order. We now delete from our list every statement which does not denote a number. We now renumber the remaining sentences calling them $n_1, n_2, \ldots n_k$. We now claim that all numbers which can be defined by finitely many words have been counted. It has been established by Cantor that the real numbers are not numberable in this fashion. He showed how to construct another number after the counting is complete. "Let p be the digit in the nth decimal place of the nth number of the set of ordered numbers. Let us form a number having 0 for its integral part and in its nth decimal place p+1 if the digit p is not 8 or 9, and 1 otherwise." The number we have just constructed is not one of the set for it differs from each of these at the nth place from the nth number. It is a number we have not counted. But the words in quotation marks above are finitely long and define this number. Therefore it is included in our enumeration. Thus our number is both included and excluded from the list, and a contradiction results.

3.3 In his original paper Richard pointed out the error in reasoning which allows the contradiction to develop. He observes that we should never have admitted our definition of a new number to be a definition. Its reference to the total set of included numbers is illegitimate, Richard says, because the set does not exist until it has been set out completely. Thus we should not say our definition should have been included in the enumeration because the definition uses a set which has no meaning in the context in which it occurs. It is improper within the enumeration to refer to this set. Thus there is no paradox. In effect the reference of the new definition is to a meaning attached to the numbering scheme which it cannot properly possess. If we adjoin out definition and its number after the enumeration is complete, it can be included without difficulty. Insert it anywhere in the sequence and increase by one the number of any statement above it.

Nagel and Newman[2] see the difficulty involved in a somewhat different light. Their paradox is constructed in a different manner. They extract a contradiction using a property "Richardian". This property refers to a statement which expresses a property not satisfied by the number assigned to the statement. Then "Richardian" is a property of natural numbers. (Nagel and Newman use number properties, not names of numbers) and hence it has a number assigned to it. When we ask if a number n is Richardian we ask if n does not have the property expressed by the statement assigned ot it. Now consider the number n which "Richardian" is assigned to.

If n is Richardian, it does have the property required, and hence it cannot be Richardian. Similarly, if n does not have this property of being Richardian, it is Richardian. Thus n is Richardian if and only if n is not Richardian, and a contradiction ensues.
Here the paradox is resolved in terms of linguistic reference. The definition of Richardian makes a reference not to any arithmetic property of a number, but to a notational property of the way the number properties were counted. The counting referred to arithmetic statements, but the numbers assigned are not simply arithmetic, and thus Richardianism should never have been admitted into the enumeration.[3]
3.4 How then do Richard and Gödel differ? Richard and Gödel both depend on a numbering of statements. But the use that each makes of the numbering scheme is different. Gödel's proof is a discussion about set membership of numbers. The recursive function theory has been developed so that decision about set membership can be made on the basis of satisfaction of certain number theorectic equations. The entire discussion can be carried out without any statement on how the set membership, in some sense, can be taken as a statement about the system itself.

Richard is not so careful. In developing his additional number, reference is made to the numbering scheme, not considered just as numbers, but as numbers representing statements in the system. The numbers are not objects in themselves for investigatioh, as they are for Gödel, but in effect are an integral part of the system in which he is operating. Thus Richard sees it as illegitimate to treat these numbers as if they could be objects for the system, since they are a part of the system. In effect these numbers carry a meaning other than themselves and which the systemic machinery is not capable of handling. When this is observed, the paradox cannot be developed.

Gödel, however, is able to use numbers as an object of discourse considered solely as numbers. The immediate object of the discussion is not the intended object, but this situation need not be recognized. As a result Gödel can be forced into no difficulties, either with the system, or with its image in the arithmetic.

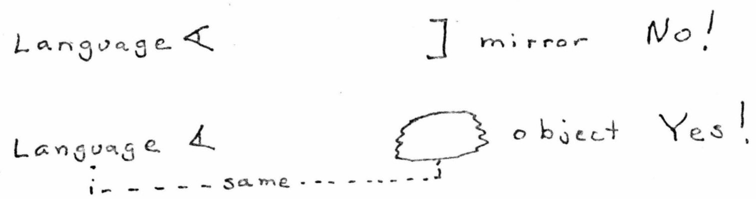
3.5 No contradiction similar to Richard's can be extracted about the system itself. The language of *Principia Mathematica*, in Gödel, never say anything explicitly about itself. It talks about relations among numbers, which are taken as its object for discussion. We are shown that the numbers model the system, but this is never proven within the system. Any attempt to force a contradiction about the system from the result uncovers the defense that the theorem is one about numbers alone, just like the

Fundamental Theorem of Arithmetic, or 2+2=4. That the system has flaws is seen but never said.

3.6 Neither can any contradiction similar to Richard's be developed from the numbers. If it could be shown that a given number both is and is not an element of a given set, we would show only that Principia Mathematica is inconsistent. Our result would still be demonstrable within the system (as of course any result would be), but the result would not be destroyed. The numbers are objects of the system and do not refer back to it. Even our contradiction would have to be recognized as such by the system. And this inconsistency could not be recognized by the system since it can make no statements about iself.

3.7 We have now seen that Gödel does not explicitly refer to what he intends in exhibiting his theorem. By using this subterfuge of indirect reference, he avoids the difficulties of the direct reference in Richard. What is the structure of this indirect reference?

We are asked to accept the proposition that objects to which a language may not refer explicitly, on pain of contradiction, it may refer to implicitly via an intermediate object. It just further be claimed that the language cannot demonstrate that indirect reference occurs, but may only invite us to see it that way. If this further claim were not made, the language could uncover its own trick and develop the contradiction anyway.



The issue appears here in the guise of self-reference. In the acceptable case the self-reference is implicit, in the unacceptable case it is explicit. The implicit self-reference is not of this kind:

(1): Sentence 2 is true.

(2): Sentence 1 is false.

This can be considered not to be implicit reference, but rather a direct reference via a transformation which would render one as:

(3): It is true that this very sentence is false.

The transformation linking references to (1) and (2) to the sentences themselves must be severed in such a way that the transformation to (3) cannot occur.

If this is done the pair become, after Gödel:

(1): Number 2 has property T.

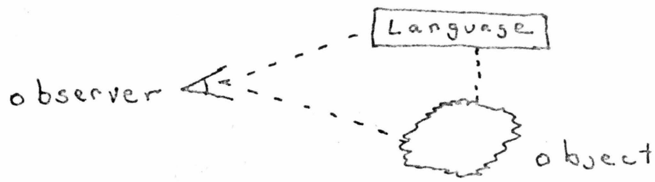
(2) Number 1 has property F.

And we might add:

(3): T and F do not both apply to the same number.

Here there is no contradiction. We may see it as a contradiction (if we number the sentences and interpret T and F in relation to the meanings of the sentences), or we may not do so. It is our choice.

It is as if there were an image of the system which is examined but which is never seen to be an image. So we have three parties to the encounter:



What the system may not say, the observer may visualize. The result of Gödel's proof must alway remain subjective since an active observer is needed for the realization of the conclusion. What is presented by Gödel is not fact about the system, but an opertunity for us to see this fact.

3.8 Our observer can intuit other things besides the system and the numbers. He also sees the world. We may see the axioms of the system as a set of transformational rules of a general type which apply to all objects of cognition. We may see physical data as new axioms for our system. Is there any way these may upset Gödel's result by their fusion in the observer with the system?

We have logic and grammar which we would consider as axioms. To this, with a status similar to Peano's natural number postulates, we would join the theories of empirical science.

The question now becomes whether Gödel's results are applicable to this larger system. In his proof on incompleteness of PM in Proposition VI, Gödel adjoins a class c of formulae which he requires only to be ω -consistent and recursive. Later he weakens recursiveness for this class to decidability, or satisfaction of his proposition V.

We shall assume that the grammar of the language we have adjoined, which would encompass modal logic at least, is ω -consistent and further that it can be stated in a recursive fashion. None of this has been demonstrated.

The question is whether the theories of empirical science, when accepted, form an ω -consistent recursive class. The issue posed cannot be demonstrated in the affirmative. Rather it will be shown that the kind of structure sought by the scientist is such a class.

ω -consistency is a rather weak condition to impose on empirical judgment. It would require that at no time do we examine every x and find it a P and at the same time conclude that not all x's are P. While we may never be in position to check all x's for P, we would certainly expect given the claim that not all x's are P that if we did check each one we would either find not all x's P or we would reject the first claim. So we can judge empirical results as ω -consistent. If we found an ω-inconsistency, we would adjust what we judged to be case in order to conform with this requirement.

What about recursiveness for these propositions? The schema we have given in the previous chapter is about recursive functions of numbers, but its intent can be extended so as to include empirical propositions. We state two requirements:

(1) Every recursive assertion can be checked by finitely many specifiable steps.

(2) Given any recursive assertion and all but one of its variables, then the remaining one can be predicted on the basis of the others by finitely many specifiable steps.

(1) is the verifiability condition. It would apply within a theoretical structure in answeringquestions within that viewpoint. (2) relates not to determing fact, but to the prediction on the basis of theory. It must be possible to clearly specify exactly what the prediction of the theory would be. This condition would surely be satisfied if such laws as science uncove rs can be committed to computers for application. If this is not possible, checking of any application of a law would be impossible since the reasoning needed could not be clearly given. It would seem any satisfactory law of science would fulfill these conditions. We would then want to claim

ω -consistent recursiveness for any satisfactory set of scientific propositions.

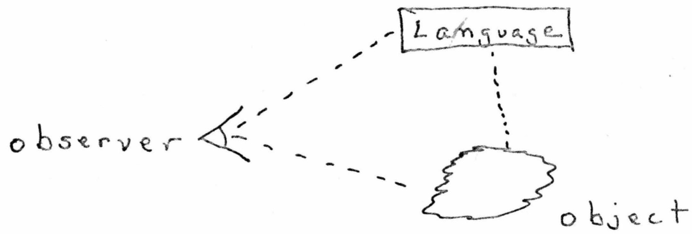
Our enlarged system would fall within the scope of Gödel's argument. It would be necessarily incomplete and powerless to demonstrate its own consistency.

We are now in position to consider language in general. We have before us in Ricard a faulty use of language and in Gödel a correct use. Chaper 4 will examine paradoxes in ordinary language to conclude that the Gödelian-Richardian difference can point to usage errors which permit the development of the paradoxes.

IV

An Extension to Ordinary Language

4.1 The concluding sections of Chaper 3 presented a structure for the reference of language necessary for Gödel's result:



What must be required of the linguistic system for this relationship to hold?

First, the language must refer outside of itself. It is not to be considered as an object for its own study, but as a means of study. In the case of Gödel's work the objects are the natural numbers. The referent could be part of the language itself, but then only if it were considered as an object and not as a language bearing meaning.

The recognition of an isomorphism between the language and its object cannot be stated in the language.

Otherwise the language would be capable of expressing, correctly, assertions about the language itself, and Richard's paradox could be developed within it. For the language in talking about its objects could express this as talk about itself, and just as it might number objects it could number iself and Richard's paradox could ensue if the final reference were to the language. So the language may not give recognition directly to itself, but may only recognize objects other than itself.

If the relation between the object and the language may not be said in the language, it must be learned some other way and for this process we have used the words show or see. What may not be said may be shown.

4.2 When we approach Gödel's work, we do possess a language in which the exact relation of our system to itself can be specified. The ordinary English, employed in Chapter 2, can express this relationship, and did so in the section on Gödel numbering. But English is a "higher level" language than Principia Mathematica since it can express more than PM. Is a higher level language essential to a Gödelian demonstration for a lower level language?

As far as simply stating the proof, such a language is not needed, and the ability to carry out such a derivation without the higher level language was needed for Proposition XI on consistency demonstrations. But while such a language can symbolically represent the proof, we have seen the reference cannot be to the language itself, and a higher level language is necessary to express the conclusion, but not to show it. But in turn the higher level language must be powerless to express comparable claims about itself. If we reject the Gödelian system of indirect reference as adequate for talk about language, then we must possess ascending series of languages each powerless to discuss itself in order to make such inquiries. The existence of such a series is doubtful.

It is questionalbe that such a series can exist, ad infinitum, in so far as we are capble of learning the language. In so far as such a language could be used, it could be learned. Some language (or native capacity) must be capable of describing the higher level language. As such it would seem our new language is nothing but a normal extension of the old to fit new circumstances. To be sure, the grammar of the new language may be more complex, but if its features can be described in English and convienient expressions for them found, it may be seen as just more English. In at least this way, many new languages can be considered as extensions of English, in much the same way as we add new nouns, or perhaps even a new verb tense. Any really different language would have to be one English is incapable of describing.

What form this might take is a puzzle.

4.3 Those who favor admitting levels to languages do so to avoid paradox.

(1) The tree is green.

(2) "The tree is green" is a sentence.

(2) would be said to be on a higher level than (1). On our view that there can only be one language, (2) would be taken as referring to a particular object, namely (1), just as (1) refers to a particular object, the green tree. No paradox can develop if the referent of "the tree is green" is something other than its meaningful occurance as (1). It may refer to itself as an object, but it may not refer to "the tree is green" as meaningful usage.

We reject any level of language theory and instead bar any reference in a language to itself as language bearing meaning. We may continue to speak of correct usage, just as well formed formulae may be picked out by standard procedures, but we will admit as proper no statement which refers to other languge directly as anything other than an event. We will not permit this reference to indicate the sentence together with its meaning.

This sentence is typed on
white paper.

The reference is acceptable since it is an object, which can be considered apart from its meaning.

This sentence is false.

The construction of this sentence is inappropriate since falseness must apply to its referent's meaning, not the simple occurence.

Using this criterion, the paradox which follows from the sentence[1]

> Every sentence is possibly either
> false or neither true nor false.

could be expected since the object of this sentence must be a sentence together with its meaning.

Gödel carefully avoids making this kind of reference and refers only to objects apart from the meaning they carry. Richard does not. He refers to numbers and then through to the meaning-sentence attached to them. Certain grammatical adjustments will be permitted, but in general language may refer only outside itself. It may not use as its object any part of the language itself, considered as language bearing meaning.

4.4 Is this criterion too strong? Do we unnecesarily exclude situations such as:

A
(1) Men are on the earth.
(2) Sentence (1) is true.

According to our reference limitation sentence (2) is illegitimate as more constructed. But the intent behind (2) can be construed as something like (1) or possibly "Don't count this answer wrong!" But (2) can usually be read in another way. Now consider:

(1) The moon is made of green cheese.

B

(2) Sentence (1) is false.

(2) can be read as saying that it is not the case that the moon is made of green cheese, i.e., the moon is not made of green cheese. This is a direct claim about the moon of the same type as (1). Cases such as A and B can be dealt with by adjusting the language involved so that it will not be forbidden.

Not all situations of such reference can be resolved in this fashion. Consider

(1) Sentence (2) is true.

C

(2) Sentence (1) is false.

What can be done with this pair? (2), according to our rendering, requires that we rewrite (1) as

(1') Sentence (2) is not true.

Or alternatively,

(1') Sentence (2) is false.

If this is to have meaning sentence (2) must be redone which requires a reconsideration of (1). This pair is illegitimate since we cannot find an equivalent pair which do not require adjustment to be acceptable.

4.5 We have examined Gödel's Theorem in relation to Richard's Paradox. We have found in Richard a double direct reference, first to a number and then to a sentence associated with that number. We have seen that this double direct reference does not occur in Gödel. We then looked at English looking at this situation and saw that either a series of language levels, or an acceptance of indirect reference was necessary to account for the meaning in Gödel. The suspicion was raised that levels of language are not possible. We examined Gödel to provoke a criterion for illegitimate reference which would not require a language level theory. It is this criterion which accounts for Gödel's success and Richard's failure. It is further suggested that the restriction of reference by language to non-linguistic events would be necessary to prevent paradox.

Anthony M. Coyne

April 1970

Appendix I: Richard's 1905 paper

The translation is from van Heijenoort's From Frege to Godel.

In its issue of 30 March 1905 the *Revue* draws attention to certain contradictions that are encountered in general set theory.

It is not necessary to go so far as the theory of ordinal numbers to find such contradictions. Here is one that presents itself the moment we study the continuum and to which some others could probably be reduced.

I am going to define a certain set of numbers, which I shall call the set E, through the following considerations.

Let us write all permutations of the twenty-six letters of the French alphabet taken two at a time, putting these permutations in alphabetical order; then, after them, all permutations taken three at a time, in alphabetical order; then, after them, all permutations taken four at a time, and so forth. These permutations may contain the same letter repeated several times; they are permutations with repetitions.

For any integer p, any permutation of the twenty-six letters taken p at a time will be in the table; and, since everything that can be written with finitely many words is a permutation of letters, everything that can be written will be in the table formed as we have just indicated.

The definition of a number being made up of words, and these words of letters, some of these permutations will be definitions of numbers. Let us cross out from our permutations all those that are not definitions of numbers.

Let u_1 be the first number defined by a permutation, u_2 the second, u_3 the third, and so on.

We thus have, written in a definite order, *all numbers that are defined by finitely many words*.

Therefore, the numbers that can be defined by finitely many words form a denumerably infinite set.

Now, here comes the contradiction. We can form a number not belonging to this set. "Let p be the digit in the nth decimal place of the nth number of the set E; let us form a number having 0 for its integral part and, in its nth decimal place, $p + 1$ if p is not 8 or 9, and 1 otherwise." This number N does not belong to the set E. If it were the nth number of the set E, the digit in its nth decimal place would be the same as the one in the nth decimal place of that number, which is not the case.

I denote by G the collection of letters between quotation marks.

The number N is defined by the words of the collection G, that is, by finitely many words; hence it should belong to the set E. But we have seen that it does not.

Such is the contradiction.

Let us show that this contradiction is only apparent. We come back to our permutations. The collection G of letters is one of these permutations; it will appear in my table. But, at the place it occupies, it has no meaning. It mentions the set E, which has not yet been defined. Hence I have to cross it out. The collection G has meaning only if the set E is totally defined, and this is not done except by infinitely many words. *Therefore there is no contradiction.*

We can make a further remark. The set containing [[the elements of]] the set E and the number N represents a new set. This new set is denumerably infinite. The number N can be inserted into the set E at a certain rank k if we increase by 1 the rank of each number of rank [[equal to or]] greater than k. Let us still denote by E the thus modified set. Then the collection of words G will define a number N′ *distinct from* N, since the number N now occupies rank k and the digit in the kth decimal place of N′ is not equal to the digit in the kth decimal place of the kth number of the set E.

Appendix II: Gödel's 1931 paper

The translation is from van Heijenoort's From Frege to Godel. The notation used in the paper differs from the exposition in Chapter 2 in that Gödel's class κ is denoted as c.

restrict the means of proof in any way). Hence a consistency proof for the system S can be carried out only by means of modes of inference that are not formalized in the system S itself, and analogous results hold for other formal systems as well, such as the Zermelo-Fraenkel axiom system of set theory.[3]

III. Theorem I can be sharpened to the effect that, even if we add finitely many axioms to the system S (or infinitely many that result from a finite number of them by "type elevation"), we do *not* obtain a complete system, provided the extended system is ω-consistent. Here a system is said to be ω-consistent if, for no property $F(x)$ of natural numbers,

$$F(1), F(2), \ldots, F(n), \ldots \text{ ad infinitum}$$

as well as

$$(Ex)\overline{F(x)}$$

are provable. (There are extensions of the system S that, while consistent, are not ω-consistent.)

IV. Theorem I still holds for all ω-consistent extensions of the system S that are obtained by the addition of *infinitely many* axioms, provided the added class of axioms is decidable [[entscheidungsdefinit]], that is, provided it is metamathematically decidable [[entscheidbar]] for every formula whether it is an axiom or not (here again we suppose that the logic used in metamathematics is that of *Principia mathematica*).

Theorems I, III, and IV can be extended also to other formal systems, for example, to the Zermelo-Fraenkel axiom system of set theory, provided the systems in question are ω-consistent.

The proofs of these theorems will appear in *Monatshefte für Mathematik und Physik*.

[3] This result, in particular, holds also for the axiom system of classical mathematics, as it has been constructed, for example, by von Neumann (*1927*).

[173]

ON FORMALLY UNDECIDABLE PROPOSITIONS OF *PRINCIPIA MATHEMATICA* AND RELATED SYSTEMS I[1]

(*1931*)

1

The development of mathematics toward greater precision has led, as is well known, to the formalization of large tracts of it, so that one can prove any theorem using nothing but a few mechanical rules. The most comprehensive formal systems that have been set up hitherto are the system of *Principia mathematica* (*PM*)[2] on the one hand and the Zermelo-Fraenkel axiom system of set theory (further developed by J. von Neumann)[3] on the other. These two systems are so comprehensive that in

[1] See a summary of the results of the present paper in *Gödel 1930b*.

[2] *Whitehead and Russell 1925*. Among the axioms of the system *PM* we include also the axiom of infinity (in this version: there are exactly denumerably many individuals), the axiom of reducibility, and the axiom of choice (for all types).

[3] See *Fraenkel 1927* and *von Neumann 1925, 1928*, and *1929*. We note that in order to complete the formalization we must add the axioms and rules of inference of the calculus of logic to the set-theoretic axioms given in the literature cited. The considerations that follow apply also to the formal systems (so far as they are available at present) constructed in recent years by Hilbert and his collaborators. See *Hilbert 1922, 1922a, 1927, Bernays 1923, von Neumann 1927*, and *Ackermann 1924*.

Note: bracketed page numbers in the margins refer to original pagination.

them all methods of proof today used in mathematics are formalized, that is, reduced to a few axioms and rules of inference. One might therefore conjecture that these axioms and rules of inference are sufficient to decide *any* mathematical question that can at all be formally expressed in these systems. It will be shown below that this is not the case, that on the contrary there are in the two systems mentioned relatively simple problems in the theory of integers[4] that cannot be decided on the basis of the axioms. This situation is not in any way due to the special nature of the systems that have been set up but holds for a wide class of formal systems; among these, in particular, are all systems that result from the two just mentioned through the addition of a finite number of axioms,[5] provided no false propositions of the kind specified in footnote 4 become provable owing to the added axioms. [174]

Before going into details, we shall first sketch the main idea of the proof, of course without any claim to complete precision. The formulas of a formal system (we restrict ourselves here to the system *PM*) in outward appearance are finite sequences of primitive signs (variables, logical constants, and parentheses or punctuation dots), and it is easy to state with complete precision *which* sequences of primitive signs are meaningful formulas and which are not.[6] Similarly, proofs, from a formal point of view, are nothing but finite sequences of formulas (with certain specifiable properties.) Of course, for metamathematical considerations it does not matter what objects are chosen as primitive signs, and we shall assign natural numbers to this use.[7] Consequently, a formula will be a finite sequence of natural numbers,[8] and a proof array a finite sequence of finite sequences of natural numbers. The metamathematical notions (propositions) thus become notions (propositions) about natural numbers or sequences of them;[9] therefore they can (at least in part) be expressed by the symbols of the system *PM* itself. In particular, it can be shown that the notions "formula", "proof array", and "provable formula" can be defined in the system *PM*; that is, we can, for example, find a formula $F(v)$ of *PM* with one free variable v (of the type of a number sequence)[10] such that $F(v)$, interpreted according to the meaning of the terms of *PM*, says: v is a provable formula. We now construct an undecidable proposition of the system *PM*, that is, a proposition A for which neither A nor *not-A* is provable, in the following manner.

[4] That is, more precisely, there are undecidable propositions in which, besides the logical constants $\overline{}$ (not), $\vee$ (or), (x) (for all), and $=$ (identical with), no other notions occur but $+$ (addition) and . (multiplication), both for natural numbers, and in which the prefixes (x), too, apply to natural numbers only.

[5] In *PM* only axioms that do not result from one another by mere change of type are counted as distinct.

[6] Here and in what follows we always understand by "formula of *PM*" a formula written without abbreviations (that is, without the use of definitions). It is well known that [in *PM*] definitions serve only to abbreviate notations and therefore are dispensable in principle.

[7] That is, we map the primitive signs one-to-one onto some natural numbers. (See how this is done on page 601.)

[8] That is, a number-theoretic function defined on an initial segment of the natural numbers. (Numbers, of course, cannot be arranged in a spatial order.)

[9] In other words, the procedure described above yields an isomorphic image of the system *PM* in the domain of arithmetic, and all metamathematical arguments can just as well be carried out in this isomorphic image. This is what we do below when we sketch the proof; that is, by "formula", "proposition", "variable", and so on, *we must always understand the corresponding objects of the isomorphic image.*

[10] It would be very easy (although somewhat cumbersome) to actually write down this formula.

[175] A formula of *PM* with exactly one free variable, that variable being of the type of the natural numbers (class of classes), will be called a *class sign*. We assume that the class signs have been arranged in a sequence in some way,[11] we denote the nth one by $R(n)$, and we observe that the notion "class sign", as well as the ordering relation R, can be defined in the system *PM*. Let α be any class sign; by $[\alpha; n]$ we denote the formula that results from the class sign α when the free variable is replaced by the sign denoting the natural number n. The ternary relation $x = [y; z]$, too, is seen to be definable in *PM*. We now define a class K of natural numbers in the following way:

$$n \,\varepsilon\, K \equiv \overline{Bew}\,[R(n); n] \tag{1}$$

(where $Bew\, x$ means: x is a provable formula).[11a] Since the notions that occur in the definiens can all be defined in *PM*, so can the notion K formed from them; that is, there is a class sign S such that the formula $[S; n]$, interpreted according to the meaning of the terms of *PM*, states that the natural number n belongs to K.[12] Since S is a class sign, it is identical with some $R(q)$; that is, we have

$$S = R(q)$$

for a certain natural number q. We now show that the proposition $[R(q); q]$ is undecidable in *PM*.[13] For let us suppose that the proposition $[R(q); q]$ were provable; then it would also be true. But in that case, according to the definitions given above, q would belong to K, that is, by (1), $\overline{Bew}\,[R(q); q]$ would hold, which contradicts the assumption. If, on the other hand, the negation of $[R(q); q]$ were provable, then $\overline{q \,\varepsilon\, K}$,[13a] that is, $Bew\,[R(q); q]$, would hold. But then $[R(q); q]$, as well as its negation, would be provable, which again is impossible.

The analogy of this argument with the Richard antinomy leaps to the eye. It is closely related to the "Liar" too;[14] for the undecidable proposition $[R(q); q]$ states that q belongs to K, that is, by (1), that $[R(q); q]$ is not provable. We therefore have before us a proposition that says about itself that it is not provable [in *PM*].[15] The
[176] method of proof just explained can clearly be applied to any formal system that, first, when interpreted as representing a system of notions and propositions, has at

[11] For example, by increasing sum of the finite sequence of integers that is the "class sign", and lexicographically for equal sums.

[11a]. The bar denotes negation.

[12] Again, there is not the slightest difficulty in actually writing down the formula S.

[13] Note that "$[R(q); q]$" (or, which means the same, "$[S; q]$") is merely a *metamathematical description* of the undecidable proposition. But, as soon as the formula S has been obtained, we can, of course, also determine the number q and, therewith, actually write down the undecidable proposition itself. [This makes no difficulty in principle. However, in order not to run into formulas of entirely unmanageable lengths and to avoid practical difficulties in the computation of the number q, the construction of the undecidable proposition would have to be slightly modified, unless the technique of abbreviation by definition used throughout in *PM* is adopted.]

[13a] [[The German text reads $\overline{n \,\varepsilon\, K}$, which is a misprint.]]

[14] Any epistemological antinomy could be used for a similar proof of the existence of undecidable propositions.

[15] Contrary to appearances, such a proposition involves no faulty circularity, for initially it [only] asserts that a certain well-defined formula (namely, the one obtained from the qth formula in the lexicographic order by a certain substitution) is unprovable. Only subsequently (and so to speak by chance) does it turn out that this formula is precisely the one by which the proposition itself was expressed.

49

its disposal sufficient means of expression to define the notions occurring in the argument above (in particular, the notion "provable formula") and in which, second, every provable formula is true in the interpretation considered. The purpose of carrying out the above proof with full precision in what follows is, among other things, to replace the second of the assumptions just mentioned by a purely formal and much weaker one.

From the remark that $[R(q); q]$ says about itself that it is not provable it follows at once that $[R(q); q]$ is true, for $[R(q); q]$ *is* indeed unprovable (being undecidable). Thus, the proposition that is undecidable *in the system PM* still was decided by metamathematical considerations. The precise analysis of this curious situation leads to surprising results concerning consistency proofs for formal systems, results that will be discussed in more detail in Section 4 (Theorem XI).

2

We now proceed to carry out with full precision the proof sketched above. First we give a precise description of the formal system P for which we intend to prove the existence of undecidable propositions. P is essentially the system obtained when the logic of PM is superposed upon the Peano axioms[16] (with the numbers as individuals and the successor relation as primitive notion).

The primitive signs of the system P are the following:

I. Constants: "$\sim$" (not), "$\vee$" (or), "Π" (for all), "0" (zero), "f" (the successor of), "(", ")" (parentheses);

II. Variables of type 1 (for individuals, that is, natural numbers including 0): "x_1", "y_1", "z_1", . . . ;

Variables of type 2 (for classes of individuals): "x_2", "y_2", "z_2", . . . ;

Variables of type 3 (for classes of classes of individuals): "x_3", "y_3", "z_3", . . . ;

And so on, for every natural number as a type.[17]

Remark: Variables for functions of two or more argument places (relations) need not be included among the primitive signs since we can define relations to be classes of ordered pairs, and ordered pairs to be classes of classes; for example, the ordered pair a, b can be defined to be $((a), (a, b))$, where (x, y) denotes the class whose sole elements are x and y, and (x) the class whose sole element is x.[18]

By a *sign of type* 1 we understand a combination of signs that has [any one of] the forms [177]

$$a, fa, ffa, fffa, \ldots, \text{ and so on},$$

where a is either 0 or a variable of type 1. In the first case, we call such a sign a *numeral*. For $n > 1$ we understand by a *sign of type n* the same thing as by a *variable of type n*. A combination of signs that has the form $a(b)$, where b is a sign of type n

[16] The addition of the Peano axioms, as well as all other modifications introduced in the system PM, merely serves to simplify the proof and is dispensable in principle.

[17] It is assumed that we have denumerably many signs at our disposal for each type of variables.

[18] Nonhomogeneous relations, too, can be defined in this manner; for example, a relation between individuals and classes can be defined to be a class of elements of the form $((x_2), ((x_1), x_2))$. Every proposition about relations that is provable in PM is provable also when treated in this manner, as is readily seen.

and a a sign of type $n + 1$, will be called an *elementary formula*. We define the class of *formulas* to be the smallest class[19] containing all elementary formulas and containing $\sim(a)$, $(a) \vee (b)$, $x\Pi(a)$ (where x may be any variable)[18a] whenever it contains a and b. We call $(a) \vee (b)$ the *disjunction* of a and b, $\sim(a)$ the *negation* and $x\Pi(a)$ a *generalization* of a. A formula in which no free variable occurs (*free variable* being defined in the well-known manner) is called a *sentential formula* [[Satzformel]]. A formula with exactly n free individual variables (and no other free variables) will be called an *n-place relation sign*; for $n = 1$ it will also be called a *class sign*.

By Subst $a\binom{v}{b}$ (where a stands for a formula, v for a variable, and b for a sign of the same type as v) we understand the formula that results from a if in a we replace v, wherever it is free, by b.[20] We say that a formula a is a *type elevation* of another formula b if a results from b when the type of each variable occurring in b is increased by the same number.

The following formulas (I–V) are called *axioms* (we write them using these abbreviations, defined in the well-known manner: $.$, $\supset$, $\equiv$, (Ex), $=$,[21] and observing the usual conventions about omitting parentheses):[22]

I. 1. $\sim(fx_1 = 0)$,
   2. $fx_1 = fy_1 \supset x_1 = y_1$,
   3. $x_2(0) . x_1\Pi(x_2(x_1) \supset x_2(fx_1)) \supset x_1\Pi(x_2(x_1))$.

[178] II. All formulas that result from the following schemata by substitution of any formulas whatsoever for p, q, r:

1. $p \vee p \supset p$, 3. $p \vee q \supset q \vee p$,
2. $p \supset p \vee q$, 4. $(p \supset q) \supset (r \vee p \supset r \vee q)$.

III. Any formula that results from either one of the two schemata

1. $v\Pi(a) \supset$ Subst $a\binom{v}{c}$,
2. $v\Pi(b \vee a) \supset b \vee v\Pi(a)$

when the following substitutions are made for a, v, b, and c (and the operation indicated by "Subst" is performed in 1):

For a any formula, for v any variable, for b any formula in which v does not occur free, and for c any sign of the same type as v, provided c does not contain any variable that is bound in a at a place where v is free.[23]

[19] Concerning this definition (and similar definitions occurring below) see *Łukasiewicz and Tarski 1930*.

[18a] Hence $x\Pi(a)$ is a formula even if x does not occur in a or is not free in a. In this case, of course, $x\Pi(a)$ means the same thing as a.

[20] In case v does not occur in a as a free variable we put Subst $a\binom{v}{b} = a$. Note that "Subst" is a metamathematical sign.

[21] $x_1 = y_1$ is to be regarded as defined by $x_2\Pi(x_2(x_1) \supset x_2(y_1))$, as in *PM* (I, ∗13) similarly for higher types).

[22] In order to obtain the axioms from the schemata listed we must therefore
(1) Eliminate the abbreviations and
(2) Add the omitted parentheses
(in II, III, and IV after carrying out the substitutions allowed).
Note that all expressions thus obtained are "formulas" in the sense specified above. (See also the exact definitions of the metamathematical notions on pp. 603–606.)

[23] Therefore c is a variable or 0 or a sign of the form $f \ldots fu$, where u is either 0 or a variable of type 1. Concerning the notion "free (bound) at a place in a", see I A 5 in *von Neumann 1927*.

IV. Every formula that results from the schema

1. $(Eu)(v\Pi(u(v) \equiv a))$

when for v we substitute any variable of type n, for u one of type $n+1$, and for a any formula that does not contain u free. This axiom plays the role of the axiom of reducibility (the comprehension axiom of set theory).

V. Every formula that results from

1. $x_1\Pi(x_2(x_1) \equiv y_2(x_1)) \supset x_2 = y_2$

by type elevation (as well as this formula itself). This axiom states that a class is completely determined by its elements.

A formula c is called an *immediate consequence* of a and b if it is the formula $(\sim(b)) \vee (c)$, and it is called an *immediate consequence* of a if it is the formula $v\Pi(a)$, where v denotes any variable. The class of *provable formulas* is defined to be the smallest class of formulas that contains the axioms and is closed under the relation "immediate consequence".[24]

We now assign natural numbers to the primitive signs of the system P by the following one-to-one correspondence: [179]

"0"	... 1	"$\sim$"	... 5	"Π"	... 9
"f"	... 3	"$\vee$"	... 7	"("	... 11
				")"	... 13;

to the variables of type n we assign the numbers of the form p^n (where p is a prime number >13). Thus we have a one-to-one correspondence by which a finite sequence of natural numbers is associated with every finite sequence of primitive signs (hence also with every formula). We now map the finite sequences of natural numbers on natural numbers (again by a one-to-one correspondence), associating the number $2^{n_1}.3^{n_2}.\dots.p_k^{n_k}$, where p_k denotes the kth prime number (in order of increasing magnitude), with the sequence $n_1, n_2, \dots, n_k$. A natural number ⟦out of a certain subset⟧ is thus assigned one-to-one not only to every primitive sign but also to every finite sequence of such signs. We denote by $\Phi(a)$ the number assigned to the primitive sign (or to the sequence of primitive signs) a. Now let some relation (or class) $R(a_1, a_2, \dots, a_n)$ between [or of] primitive signs or sequences of primitive signs be given. With it we associate the relation (or class) $R'(x_1, x_2, \dots, x_n)$ between [or of] natural numbers that obtains between $x_1, x_2, \dots, x_n$ if and only if there are some $a_1, a_2, \dots, a_n$ such that $x_i = \Phi(a_i)$ $(i = 1, 2, \dots, n)$ and $R(a_1, a_2, \dots, a_n)$ hold. The relations between (or classes of) natural numbers that in this manner are associated with the metamathematical notions defined so far, for example, "variable", "formula", "sentential formula", "axiom", "provable formula", and so on, will be denoted by the same words in SMALL CAPITALS. The proposition that there are undecidable problems in the system P, for example, reads thus: There are SENTENTIAL FORMULAS a such that neither a nor the NEGATION of a is a PROVABLE FORMULA.

We now insert a parenthetic consideration that for the present has nothing to do

[24] The rule of substitution is rendered superfluous by the fact that all possible substitutions have already been carried out in the axioms themselves. (This procedure was used also by *von Neumann 1927*.)

52

with the formal system P. First we give the following definition: A number-theoretic function[25] $\varphi(x_1, x_2, \ldots, x_n)$ is said to be *recursively defined in terms of* the number-theoretic functions $\psi(x_1, x_2, \ldots, x_{n-1})$ and $\mu(x_1, x_2, \ldots, x_{n+1})$ if

$$\begin{aligned} \varphi(0, x_2, \ldots, x_n) &= \psi(x_2, \ldots, x_n), \\ \varphi(k+1, x_2, \ldots, x_n) &= \mu(k, \varphi(k, x_2, \ldots, x_n), x_2, \ldots, x_n) \end{aligned} \tag{2}$$

hold for all $x_2, \ldots, x_n, k$.[26]

[180] A number-theoretic function φ is said to be *recursive* if there is a finite sequence of number-theoretic functions $\varphi_1, \varphi_2, \ldots, \varphi_n$ that ends with φ and has the property that every function φ_k of the sequence is recursively defined in terms of two of the preceding functions, or results from any of the preceding functions by substitution,[27] or, finally, is a constant or the successor function $x + 1$. The length of the shortest sequence of φ_i corresponding to a recursive function φ is called its *degree*. A relation $R(x_1, \ldots, x_n)$ between natural numbers is said to be recursive[28] if there is a recursive function $\varphi(x_1, \ldots, x_n)$ such that, for all $x_1, x_2, \ldots, x_n$,

$$R(x_1, \ldots, x_n) \sim [\varphi(x_1, \ldots, x_n) = 0].^{29}$$

The following theorems hold:

I. *Every function (relation) obtained from recursive functions (relations) by substitution of recursive functions for the variables is recursive; so is every function obtained from recursive functions by recursive definition according to schema (2);*

II. *If R and S are recursive relations, so are $\overline{R}$ and $R \vee S$ (hence also R & S);*

III. *If the functions $\varphi(\mathfrak{x})$ and $\psi(\mathfrak{y})$ are recursive, so is the relation $\varphi(\mathfrak{x}) = \psi(\mathfrak{y})$;*[30]

IV. *If the function $\varphi(\mathfrak{x})$ and the relation $R(x, \mathfrak{y})$ are recursive, so are the relations S and T defined by*

$$S(\mathfrak{x}, \mathfrak{y}) \sim (Ex)[x \leqq \varphi(\mathfrak{x}) \ \&\ R(x, \mathfrak{y})]$$

and

$$T(\mathfrak{x}, \mathfrak{y}) \sim (x)[x \leqq \varphi(\mathfrak{x}) \rightarrow R(x, \mathfrak{y})],$$

as well as the function ψ defined by

$$\psi(\mathfrak{x}, \mathfrak{y}) = \varepsilon x[x \leqq \varphi(\mathfrak{x}) \ \&\ R(x, \mathfrak{y})],$$

where $\varepsilon x F(x)$ means the least number x for which $F(x)$ holds and 0 in case there is no such number.

[25] That is, its domain of definition is the class of nonnegative integers (or of n-tuples of nonnegative integers) and its values are nonnegative integers.

[26] In what follows, lower-case italic letters (with or without subscripts) are always variables for nonnegative integers (unless the contrary is expressly noted).

[27] More precisely, by substitution of some of the preceding functions at the argument places of one of the preceding functions, for example, $\varphi_k(x_1, x_2) = \varphi_p[\varphi_q(x_1, x_2), \varphi_r(x_2)]$ $(p, q, r < k)$. Not all variables on the left side need occur on the right side (the same applies to the recursion schema (2)).

[28] We include classes among relations (as one-place relations). Recursive relations R, of course, have the property that for every given n-tuple of numbers it can be decided whether $R(x_1, \ldots, x_n)$ holds or not.

[29] Whenever formulas are used to express a meaning (in particular, in all formulas expressing metamathematical propositions or notions), Hilbert's symbolism is employed. See *Hilbert and Ackermann 1928*.

[30] We use German letters, $\mathfrak{x}$, $\mathfrak{y}$, as abbreviations for arbitrary n-tuples of variables, for example $x_1, x_2, \ldots, x_n$.

Theorem I follows at once from the definition of "recursive". Theorems II and III are consequences of the fact that the number-theoretic functions

$$\alpha(x), \quad \beta(x, y), \quad \gamma(x, y),$$

corresponding to the logical notions $\overline{}$, $\vee$, and $=$, namely,

$$\alpha(0) = 1, \ \alpha(x) = 0 \quad \text{for } x \neq 0,$$

$$\beta(0, x) = \beta(x, 0) = 0, \quad \beta(x, y) = 1 \quad \text{when } x \text{ and } y \text{ are both } \neq 0,$$

$$\gamma(x, y) = 0 \quad \text{when } x = y, \qquad \gamma(x, y) = 1 \quad \text{when } x \neq y,$$ [181]

are recursive, as we can readily see. The proof of Theorem IV is briefly as follows. By assumption there is a recursive $\rho(x, \mathfrak{y})$ such that

$$R(x, \mathfrak{y}) \sim [\rho(x, \mathfrak{y}) = 0].$$

We now define a function $\chi(x, \mathfrak{y})$ by the recursion schema (2) in the following way:

$$\chi(0, \mathfrak{y}) = 0,$$

$$\chi(n + 1, \mathfrak{y}) = (n + 1).a + \chi(n, \mathfrak{y}).\alpha(a),$$ [31]

where $a = \alpha[\alpha(\rho(0, \mathfrak{y}))].\alpha[\rho(n + 1, \mathfrak{y})].\alpha[\chi(n, \mathfrak{y})]$. Therefore $\chi(n + 1, \mathfrak{y})$ is equal either to $n + 1$ (if $a = 1$) or to $\chi(n, \mathfrak{y})$ (if $a = 0$).[32] The first case clearly occurs if and only if all factors of a are 1, that is, if

$$\overline{R}(0, \mathfrak{y}) \ \& \ R(n + 1, \mathfrak{y}) \ \& \ [\chi(n, \mathfrak{y}) = 0]$$

holds. From this it follows that the function $\chi(n, \mathfrak{y})$ (considered as a function of n) remains 0 up to [[but not including]] the least value of n for which $R(n, \mathfrak{y})$ holds and, from there on, is equal to that value. (Hence, in case $R(0, \mathfrak{y})$ holds, $\chi(n, \mathfrak{y})$ is constant and equal to 0.) We have, therefore,

$$\psi(\mathfrak{x}, \mathfrak{y}) = \chi(\varphi(\mathfrak{x}), \mathfrak{y}),$$

$$S(\mathfrak{x}, \mathfrak{y}) \sim R[\psi(\mathfrak{x}, \mathfrak{y}), \mathfrak{y}].$$

The relation T can, by negation, be reduced to a case analogous to that of S. Theorem IV is thus proved.

The functions $x + y$, $x.y$, and x^y, as well as the relations $x < y$ and $x = y$, are recursive, as we can readily see. Starting from these notions, we now define a number of functions (relations) 1–45, each of which is defined in terms of preceding ones by the procedures given in Theorems I–IV. In most of these definitions several of the steps allowed by Theorems I–IV are condensed into one. Each of the functions (relations) 1–45, among which occur, for example, the notions "FORMULA", "AXIOM", and "IMMEDIATE CONSEQUENCE", is therefore recursive. [182]

1. $x/y \equiv (Ex)[z \leqq x \ \& \ x = y.z]$,[33]
x is divisible by y.[34]

[31] We assume familiarity with the fact that the functions $x + y$ (addition) and $x \,.\, y$ (multiplication) are recursive.

[32] a cannot take values other than 0 and 1, as can be seen from the definition of α.

[33] The sign $\equiv$ is used in the sense of "equality by definition"; hence in definitions it stands for either $=$ or $\sim$ (otherwise, the symbolism is Hilbert's).

[34] Wherever one of the signs (x), (Ex), or εx occurs in the definitions below, it is followed by a bound on x. This bound merely serves to ensure that the notion defined is recursive (see Theorem IV). But in most cases the *extension* of the notion defined would not change if this bound were omitted.

54

2. $\mathrm{Prim}(x) \equiv \overline{(Ex)}[z \leqq x \ \&\ z \neq 1 \ \&\ z \neq x \ \&\ x/z] \ \&\ x > 1$,
x is a prime number.

3. $0 \, Pr \, x \equiv 0$,
$(n + 1) \, Pr \, x \equiv \varepsilon y[y \leqq x \ \&\ \mathrm{Prim}(y) \ \&\ x/y \ \&\ y > n \, Pr \, x]$,
$n \, Pr \, x$ is the nth prime number (in order of increasing magnitude) contained in x.[34a]

4. $0! \equiv 1$,
$(n + 1)! \equiv (n + 1) . n!$.

5. $Pr(0) \equiv 0$,
$Pr(n + 1) \equiv \varepsilon y[y \leqq \{Pr(n)\}! + 1 \ \&\ \mathrm{Prim}(y) \ \&\ y > Pr(n)]$,
$Pr(n)$ is the nth prime number (in order of increasing magnitude).

6. $n \, Gl \, x \equiv \varepsilon y[y \leqq x \ \&\ x/(n \, Pr \, x)^y \ \&\ \overline{x/(n \, Pr \, x)^{y+1}}]$,
$n \, Gl \, x$ is the nth term of the number sequence assigned to the number x (for $n > 0$ and n not greater than the length of this sequence).

7. $l(x) \equiv \varepsilon y[y \leqq x \ \&\ y \, Pr \, x > 0 \ \&\ (y + 1) \, Pr \, x = 0]$,
$l(x)$ is the length of the number sequence assigned to x.

8. $x*y \equiv \varepsilon z\{z \leqq [Pr(l(x) + l(y))]^{x+y} \ \&\ (n)[n \leqq l(x) \rightarrow n \, Gl \, z = n \, Gl \, x] \ \&$
$(n)[0 < n \leqq l(y) \rightarrow (n + l(x)) \, Gl \, z = n \, Gl \, y]\}$,
$x*y$ corresponds to the operation of "concatenating" two finite number sequences.

9. $R(x) \equiv 2^x$,
$R(x)$ corresponds to the number sequence consisting of x alone (for $x > 0$).

10. $E(x) \equiv R(11)*x*R(13)$,
$E(x)$ corresponds to the operation of "enclosing within parentheses" (11 and 13 are assigned to the primitive signs "(" and ")", respectively).

11. $n \, \mathrm{Var} \, x \equiv (Ez)[13 < z \leqq x \ \&\ \mathrm{Prim}(z) \ \&\ x = z^n] \ \&\ n \neq 0$,
x is a VARIABLE OF TYPE n.

12. $\mathrm{Var}(x) \equiv (En)[n \leqq x \ \&\ n \, \mathrm{Var} \, x]$,
x is a VARIABLE.

13. $\mathrm{Neg}(x) \equiv R(5)*E(x)$,
$\mathrm{Neg}(x)$ is the NEGATION of x.

[183] 14. $x \, \mathrm{Dis} \, y \equiv E(x)*R(7)*E(y)$,
x Dis y is the DISJUNCTION of x and y.

15. $x \, \mathrm{Gen} \, y \equiv R(x)*R(9)*E(y)$,
x Gen y is the GENERALIZATION of y with respect to the VARIABLE x (provided x is a VARIABLE).

16. $0 \, N \, x \equiv x$,
$(n + 1) \, N \, x \equiv R(3)*n \, N \, x$,
$n \, N \, x$ corresponds to the operation of "putting the sign 'f' n times in front of x".

17. $Z(n) \equiv n \, N \, [R(1)]$,
$Z(n)$ is the NUMERAL denoting the number n.

18. $\mathrm{Typ}_1'(x) \equiv (Em, n)\{m, n \leqq x \ \&\ [m = 1 \vee 1 \, \mathrm{Var} \, m] \ \&\ x = n \, N \, [R(m)]\}$,[34b]
x is a SIGN OF TYPE 1.

[34a] For $0 < n \leqq z$, where z is the number of distinct prime factors of x. Note that $n \, Pr \, x = 0$ for $n = z + 1$.

[34b] $m, n \leqq x$ stands for $m \leqq x \ \&\ n \leqq x$ (similarly for more than two variables).

19. $\mathrm{Typ}_n(x) \equiv [n = 1 \ \& \ \mathrm{Typ}_1'(x)] \vee [n > 1 \ \&$
$(Ev)\{v \leqq x \ \& \ n \ \mathrm{Var}\ v \ \& \ x = R(v)\}]$,

x is a SIGN OF TYPE n.

20. $Elf(x) \equiv (Ey, z, n)[y, z, n \leqq x \ \& \ \mathrm{Typ}_n(y) \ \&$
$\mathrm{Typ}_{n+1}(z) \ \& \ x = z*E(y)]$,

x is an ELEMENTARY FORMULA.

21. $Op(x, y, z) \equiv x = \mathrm{Neg}(y) \vee x = y \ \mathrm{Dis}\ z \vee (Ev)[v \leqq x \ \& \ \mathrm{Var}(v) \ \&$
$x = v \ \mathrm{Gen}\ y]$.

22. $FR(x) \equiv (n)\{0 < n \leqq l(x) \rightarrow Elf(n \ Gl \ x) \vee (Ep, q)[0 < p, q < n \ \&$
$Op(n \ Gl \ x, p \ Gl \ x, q \ Gl \ x)]\} \ \& \ l(x) > 0$,

x is a SEQUENCE OF FORMULAS, each of which either is an ELEMENTARY FORMULA or results from the preceding FORMULAS through the operations of NEGATION, DISJUNCTION, or GENERALIZATION.

23. $\mathrm{Form}(x) \equiv (En)\{n \leqq (Pr[l(x)^2])^{x \cdot [l(x)]^2} \ \& \ FR(n) \ \& \ x = [l(n)] \ Gl \ n\}$,[35]

x is a FORMULA (that is, the last term of a FORMULA SEQUENCE n).

24. $v \ \mathrm{Geb}\ n, x \equiv \mathrm{Var}(v) \ \& \ \mathrm{Form}(x) \ \& \ (Ea, b, c)[a, b, c \leqq x \ \&$
$x = a*(v \ \mathrm{Gen}\ b)*c \ \& \ \mathrm{Form}(b) \ \& \ l(a) + 1 \leqq n \leqq l(a) + l(v \ \mathrm{Gen}\ b)]$,

the VARIABLE v is BOUND in x at the nth place.

25. $v \ Fr \ n, x \equiv \mathrm{Var}(v) \ \& \ \mathrm{Form}(x) \ \& \ v = n \ Gl \ x \ \& \ n \leqq l(x) \ \& \ \overline{v \ \mathrm{Geb}\ n, x}$, [184]

the VARIABLE v is FREE in x at the nth place.

26. $v \ Fr \ x \equiv (En)[n \leqq l(x) \ \& \ v \ Fr \ n, x]$,

v occurs as a FREE VARIABLE in x.

27. $Su \ x\binom{n}{y} \equiv \varepsilon z\{z \leqq [Pr(l(x) + l(y))]^{x+y} \ \& \ [(Eu, v) \ u, v \leqq x \ \&$
$x = u*R(n \ Gl \ x)*v \ \& \ z = u*y*v \ \& \ n = l(u) + 1]\}$,

$Su \ x\binom{n}{y}$ results from x when we substitute y for the nth term of x (provided that $0 < n \leqq l(x)$).

28. $0 \ St \ v, x \equiv \varepsilon n\{n \leqq l(x) \ \& \ v \ Fr \ n, x \ \& \ \overline{(Ep)}[n < p \leqq l(x) \ \& \ v \ Fr \ p, x]\}$,
$(k + 1) \ St \ v, x \equiv \varepsilon n\{n < k \ St \ v, x \ \& \ v \ Fr \ n, x \ \& \ \overline{(Ep)}[n < p < k \ St \ v, x$
$\& \ v \ Fr \ p, x]\}$,

$k \ St \ v, x$ is the $(k + 1)$th place in x (counted from the right end of the FORMULA x) at which v is FREE in x (and 0 in case there is no such place).

29. $A(v, x) \equiv \varepsilon n\{n \leqq l(x) \ \& \ n \ St \ v, x = 0\}$,

$A(v, x)$ is the number of places at which v is FREE in x.

30. $Sb_0(x_y^v) \equiv x$,
$Sb_{k+1}(x_y^v) \equiv Su \ [Sb_k(x_y^v)]\binom{k \ St \ v, x}{y}$.

31. $Sb(x_y^v) \equiv Sb_{A(v,x)}(x_y^v)$,[36]

$Sb(x_y^v)$ is the notion SUBST $a\binom{v}{b}$ defined above.[37]

32. $x \ \mathrm{Imp}\ y \equiv [\mathrm{Neg}(x)] \ \mathrm{Dis}\ y$,
$x \ \mathrm{Con}\ y \equiv \mathrm{Neg}\{[\mathrm{Neg}(x)] \ \mathrm{Dis}\ [\mathrm{Neg}(y)]\}$,

[35] That $n \leqq (Pr([l(x)]^2))^{x \cdot [l(x)]^2}$ provides a bound can be seen thus: The length of the shortest sequence of formulas that corresponds to x can at most be equal to the number of subformulas of x. But there are at most $l(x)$ subformulas of length 1, at most $l(x) - 1$ of length 2, and so on, hence altogether at most $l(x)(l(x) + 1)/2 \leqq [l(x)]^2$. Therefore all prime factors of n can be assumed to be less than $Pr([l(x)]^2)$, their number $\leqq [(lx)]^2$, and their exponents (which are subformulas of x) $\leqq x$.

[36] In case v is not a VARIABLE or x is not a FORMULA, $Sb(x_y^v) = x$.

[37] Instead of $Sb[Sb(x_y^v)_z^w]$ we write $Sb(x_{y\ z}^{v\ w})$ (and similarly for more than two VARIABLES).

$x \text{ Aeq } y \equiv (x \text{ Imp } y) \text{ Con } (y \text{ Imp } x)$,
$v \text{ Ex } y \equiv \text{Neg}\{v \text{ Gen } [\text{Neg}(y)]\}$.

33. $n\ Th\ x \equiv \varepsilon y\{y \leq x^{(x^n)} \ \&\ (k)[k \leq l(x) \rightarrow (k\ Gl\ x \leq 13 \ \&\ k\ Gl\ y = k\ Gl\ x) \vee (k\ Gl\ x > 13 \ \&\ k\ Gl\ y = k\ Gl\ x . [1\ Pr\ (k\ Gl\ x)]^n)]\}$,

$n\ Th\ x$ is the nth TYPE ELEVATION of x (in case x and $n\ Th\ x$ are FORMULAS).

Three specific numbers, which we denote by z_1, z_2, and z_3, correspond to the Axioms I, 1–3, and we define

34. $Z\text{-}Ax(x) \equiv (x = z_1 \vee x = z_2 \vee x = z_3)$.

[185] 35. $A_1\text{-}Ax(x) \equiv (Ey)[y \leq x \ \&\ \text{Form}(y) \ \&\ x = (y \text{ Dis } y) \text{ Imp } y]$,

x is a FORMULA resulting from Axiom schema II, 1 by substitution. Analogously, $A_2\text{-}Ax$, $A_3\text{-}Ax$, and $A_4\text{-}Ax$ are defined for Axioms [rather, Axiom Schemata] II, 2–4.

36. $A\text{-}Ax(x) \equiv A_1\text{-}Ax(x) \vee A_2\text{-}Ax(x) \vee A_3\text{-}Ax(x) \vee A_4\text{-}Ax(x)$,

x is a FORMULA resulting from a propositional axiom by substitution.

37. $Q(z, y, v) \equiv \overline{(En, m, w)}[n \leq l(y) \ \&\ m \leq l(z) \ \&\ w \leq z \ \&\ w = m\ Gl\ z \ \&\ w \text{ Geb } n, y \ \&\ v\ Fr\ n, y]$

z does not contain any VARIABLE BOUND in y at a place at which v is FREE.

38. $L_1\text{-}Ax(x) \equiv (Ev, y, z, n)\{v, y, z, n \leq x \ \&\ n \text{ Var } v \ \&\ \text{Typ}_n(z) \ \&\ \text{Form}(y) \ \&\ Q(z, y, v) \ \&\ x = (v \text{ Gen } y) \text{ Imp } [Sb(y_z^v)]\}$,

x is a FORMULA resulting from Axiom schema III, 1 by substitution.

39. $L_2\text{-}Ax(x) \equiv (Ev, q, p)\{v, q, p \leq x \ \&\ \text{Var}(v) \ \&\ \text{Form}(p) \ \&\ \overline{v\ Fr\ p} \ \&\ \text{Form}(q) \ \&\ x = [v \text{ Gen } (p \text{ Dis } q)] \text{ Imp } [p \text{ Dis } (v \text{ Gen } q)]\}$,

x is a FORMULA resulting from Axiom schema III, 2 by substitution.

40. $R\text{-}Ax(x) \equiv (Eu, v, y, n)[u, v, y, n \leq x \ \&\ n \text{ Var } v \ \&\ (n+1) \text{ Var } u \ \&\ \overline{u\ Fr\ y} \ \&\ \text{Form}(y) \ \&\ x = u \text{ Ex } \{v \text{ Gen } [[R(u) * E(R(v))] \text{ Aeq } y]\}]$,

x is a FORMULA resulting from Axiom schema IV, 1 by substitution.

A specific number z_4 corresponds to Axiom V, 1, and we define:

41. $M\text{-}Ax(x) \equiv (En)[n \leq x \ \&\ x = n\ Th\ z_4]$.

42. $Ax(x) \equiv Z\text{-}Ax(x) \vee A\text{-}Ax(x) \vee L_1\text{-}Ax(x) \vee L_2\text{-}Ax(x) \vee R\text{-}Ax(x) \vee M\text{-}Ax(x)$,

x is an AXIOM.

43. $Fl(x, y, z) \equiv y = z \text{ Imp } x \vee (Ev)[v \leq x \ \&\ \text{Var}(v) \ \&\ x = v \text{ Gen } y]$,

x is an IMMEDIATE CONSEQUENCE of y and z.

[186] 44. $Bw(x) \equiv (n)\{0 < n \leq l(x) \rightarrow Ax(n\ Gl\ x) \vee (Ep, q)[0 < p, q < n \ \&\ Fl(n\ Gl\ x, p\ Gl\ x, q\ Gl\ x)]\} \ \&\ l(x) > 0$,

x is a PROOF ARRAY (a finite sequence of FORMULAS, each of which is either an AXIOM or an IMMEDIATE CONSEQUENCE of two of the preceding FORMULAS.

45. $x\ B\ y \equiv Bw(x) \ \&\ [l(x)]\ Gl\ x = y$,

x is a PROOF of the FORMULA y.

46. $\text{Bew}(x) \equiv (Ey)y\ B\ x$,

x is a PROVABLE FORMULA. ($\text{Bew}(x)$ is the only one of the notions 1–46 of which we cannot assert that it is recursive.)

The fact that can be formulated vaguely by saying: every recursive relation is definable in the system P (if the usual meaning is given to the formulas of this system), is expressed in precise language, *without* reference to any interpretation of the formulas of P, by the following theorem:

Theorem V. *For every recursive relation* $R(x_1, \ldots, x_n)$ *there exists an* n-*place*

RELATION SIGN r *(with the* FREE VARIABLES[38] $u_1, u_2, \ldots, u_n$) *such that for all n-tuples of numbers* $(x_1, \ldots, x_n)$ *we have*

$$R(x_1, \ldots, x_n) \rightarrow \text{Bew}[Sb(r^{u_1 \ldots u_n}_{Z(x_1)\ldots Z(x_n)})], \tag{3}$$

$$\bar{R}(x_1, \ldots, x_n) \rightarrow \text{Bew}[\text{Neg}(Sb(r^{u_1 \ldots u_n}_{Z(x_1)\ldots Z(x_n)}))]. \tag{4}$$

We shall give only an outline of the proof of this theorem because the proof does not present any difficulty in principle and is rather long.[39] We prove the theorem for all relations $R(x_1, \ldots, x_n)$ of the form $x_1 = \varphi(x_2, \ldots, x_n)$[40] (where φ is a recursive function) and we use induction on the degree of φ. For functions of degree 1 (that is, constants and the function $x + 1$) the theorem is trivial. Assume now that φ is of degree m. It results from functions of lower degrees, $\varphi_1, \ldots, \varphi_k$, through the operations of substitution or recursive definition. Since by the induction hypothesis everything has already been proved for $\varphi_1, \ldots, \varphi_k$, there are corresponding RELATION SIGNS, $r_1, \ldots, r_k$, such that (3) and (4) hold. The processes of definition by which φ results from $\varphi_1, \ldots, \varphi_k$ (substitution and recursive definition) can both be formally reproduced in the system P. If this is done, a new RELATION SIGN r is obtained from $r_1, \ldots, r_k$,[41] and, using the induction hypothesis, we can prove without difficulty that (3) and (4) hold for it. A RELATION SIGN r assigned to a recursive relation[42] by this procedure will be said to be recursive. [187]

We now come to the goal of our discussions. Let κ be any class of FORMULAS. We denote by $\text{Flg}(\kappa)$ (the set of consequences of κ) the smallest set of FORMULAS that contains all FORMULAS of κ and all AXIOMS and is closed under the relation "IMMEDIATE CONSEQUENCE". κ is said to be ω-consistent if there is no CLASS SIGN a such that

$$(n)[Sb(a^{v}_{Z(n)})\ \varepsilon\ \text{Flg}(\kappa)]\ \&\ [\text{Neg}(v\ \text{Gen}\ a)]\ \varepsilon\ \text{Flg}(\kappa),$$

where v is the FREE VARIABLE of the CLASS SIGN a.

Every ω-consistent system, of course, is consistent. As will be shown later, however, the converse does not hold.

The general result about the existence of undecidable propositions reads as follows:

Theorem VI. *For every ω-consistent recursive class κ of* FORMULAS *there are recursive* CLASS SIGNS r *such that neither* v Gen r *nor* Neg(v Gen r) *belongs to* $\text{Flg}(\kappa)$ (where v is the FREE VARIABLE of r).

Proof. Let κ be any recursive ω-consistent class of FORMULAS. We define

$$Bw_\kappa(x) \equiv (n)[n \leqq l(x) \rightarrow Ax(n\ Gl\ x) \vee (n\ Gl\ x)\ \varepsilon\ \kappa\ \vee$$
$$(Ep, q)\{0 < p, q < n\ \&\ Fl(n\ Gl\ x, p\ Gl\ x, q\ Gl\ x)\}]\ \&\ l(x) > 0 \tag{5}$$

[38] The VARIABLES $u_1, \ldots, u_n$ can be chosen arbitrarily. For example, there always is an r with the FREE VARIABLES 17, 19, 23, ..., and so on, for which (3) and (4) hold.

[39] Theorem V, of course, is a consequence of the fact that in the case of a recursive relation R it can, for every n-tuple of numbers, be decided *on the basis of the axioms of the system P* whether the relation R obtains or not.

[40] From this it follows at once that the theorem holds for every recursive relation, since any such relation is equivalent to $0 = \varphi(x_1, \ldots, x_n)$, where φ is recursive.

[41] When this proof is carried out in detail, r, of course, is not defined indirectly with the help of its meaning but in terms of its purely formal structure.

[42] Which, therefore, in the usual interpretation expresses the fact that this relation holds.

58

(see the analogous notion 44),

$$x \; B_\kappa \; y \equiv Bw_\kappa(x) \; \& \; [l(x)] \; Gl \; x = \tag{6}$$

$$\text{Bew}_\kappa(x) \equiv (Ey) y \; B_\kappa \; x \tag{6.1}$$

(see the analogous notions 45 and 46).

We obviously have

$$(x)[\text{Bew}_\kappa(x) \sim x \; \varepsilon \; \text{Flg}(\kappa)] \tag{7}$$

and

$$(x)[\text{Bew}(x) \rightarrow \text{Bew}_\kappa(x)]. \tag{8}$$

[188] We now define the relation

$$Q(x, y) \equiv \overline{x \; B_\kappa \; [Sb(y^{19}_{Z(y)})]}. \tag{8.1}$$

Since $x \; B_\kappa \; y$ (by (6) and (5)) and $Sb(y^{19}_{Z(y)})$ (by Definitions 17 and 31) are recursive, so is $Q(x, y)$. Therefore, by Theorem V and (8) there is a RELATION SIGN q (with the FREE VARIABLES 17 and 19) such that

$$\overline{x \; B_\kappa \; [Sb(y^{19}_{Z(y)})]} \rightarrow \text{Bew}_\kappa[Sb(q^{17}_{Z(x)}{}^{19}_{Z(y)})], \tag{9}$$

and

$$x \; B_\kappa \; [Sb(y^{19}_{Z(y)})] \rightarrow \text{Bew}_\kappa[\text{Neg}(Sb(q^{17}_{Z(x)}{}^{19}_{Z(y)}))]. \tag{10}$$

We put

$$p = 17 \text{ Gen } q \tag{11}$$

(p is a CLASS SIGN with the FREE VARIABLE 19) and

$$r = Sb(q^{19}_{Z(p)}) \tag{12}$$

(r is a recursive CLASS SIGN[43] with the FREE VARIABLE 17).

Then we have

$$Sb(p^{19}_{Z(p)}) = Sb([17 \text{ Gen } q]^{19}_{Z(p)}) = 17 \text{ Gen } Sb(q^{19}_{Z(p)}) = 17 \text{ Gen } r \tag{13}$$

(by (11) and (12));[44] furthermore

$$Sb(q^{17}_{Z(x)}{}^{19}_{Z(p)}) = Sb(r^{17}_{Z(x)}) \tag{14}$$

(by (12)). If we now substitute p for y in (9) and (10) and take (13) and (14) into account, we obtain

$$\overline{x \; B_\kappa \; (17 \text{ Gen } r)} \rightarrow \text{Bew}_\kappa[Sb(r^{17}_{Z(x)})], \tag{15}$$

$$x \; B_\kappa \; (17 \text{ Gen } r) \rightarrow \text{Bew}_\kappa[Neg(Sb(r^{17}_{Z(x)}))]. \tag{16}$$

[189] This yields:

1. 17 Gen r is not κ-PROVABLE.[45] For, if it were, there would (by (6.1)) be an n such

[43] Since r is obtained from the recursive RELATION SIGN q through the replacement of a VARIABLE by a definite number, p. [Precisely stated the final part of this footnote (which refers to a side remark unnecessary for the proof) would read thus: "REPLACEMENT of a VARIABLE by the NUMERAL for p."]

[44] The operations Gen and Sb, of course, can always be interchanged in case they refer to different VARIABLES.

[45] By "x is κ-provable" we mean $x \; \varepsilon \; \text{Flg}(\kappa)$, which, by (7), means the same thing as $\text{Bew}_\kappa(x)$.

59

that $n\, B_\kappa$ (17 Gen r). Hence by (16) we would have $\text{Bew}_\kappa[\text{Neg}(Sb(r^{17}_{Z(n)}))]$, while, on the other hand, from the κ-PROVABILITY of 17 Gen r that of $Sb(r^{17}_{Z(n)})$ follows. Hence, κ would be inconsistent (and a fortiori ω-inconsistent).

2. Neg(17 Gen r) is not κ-PROVABLE. Proof: As has just been proved, 17 Gen r is not κ-PROVABLE; that is (by (6.1)), $(n)\overline{n\, B_\kappa\,(17 \text{ Gen } r)}$ holds. From this, $(n)\text{Bew}_\kappa[Sb(r^{17}_{Z(n)})]$ follows by (15), and that, in conjunction with Bew_κ[Neg(17 Gen r)], is incompatible with the ω-consistency of κ.

17 Gen r is therefore undecidable on the basis of κ, which proves Theorem VI.

We can readily see that the proof just given is constructive;[45a] that is, the following has been proved in an intuitionistically unobjectionable manner: Let an arbitrary recursively defined class κ of FORMULAS be given. Then, if a formal decision (on the basis of κ) of the SENTENTIAL FORMULA 17 Gen r (which [for each κ] can actually be exhibited) is presented to us, we can actually give

1. A PROOF of Neg(17 Gen r);
2. For any given n, a PROOF of $Sb(r^{17}_{Z(n)})$.

That is, a formal decision of 17 Gen r would have the consequence that we could actually exhibit an ω-inconsistency.

We shall say that a relation between (or a class of) natural numbers $R(x_1, \ldots, x_n)$ is *decidable* [[*entscheidungsdefinit*]] if there exists an n-place RELATION SIGN r such that (3) and (4) (see Theorem V) hold. In particular, therefore, by Theorem V every recursive relation is decidable. Similarly, a RELATION SIGN will be said to be *decidable* if it corresponds in this way to a decidable relation. Now it suffices for the existence of undecidable propositions that the class κ be ω-consistent and decidable. For the decidability carries over from κ to $x\, B_\kappa\, y$ (see (5) and (6)) and to $Q(x, y)$ (see (8.1)), and only this was used in the proof given above. In this case the undecidable proposition has the form v Gen r, where r is a decidable CLASS SIGN. (Note that it even suffices that κ be decidable in the system enlarged by κ.) [190]

If, instead of assuming that κ is ω-consistent, we assume only that it is consistent, then, although the existence of an undecidable proposition does not follow [by the argument given above], it does follow that there exists a property (r) for which it is possible neither to give a counterexample nor to prove that it holds of all numbers. For in the proof that 17 Gen r is not κ-PROVABLE only the consistency of κ was used (see p. 608). Moreover from $\overline{\text{Bew}_\kappa}$(17 Gen r) it follows by (15) that, for every number x, $Sb(r^{17}_{Z(x)})$ is κ-PROVABLE and consequently that $\text{Neg}(Sb(r^{17}_{Z(x)}))$ is not κ-PROVABLE for any number.

If we adjoin Neg(17 Gen r) to κ, we obtain a class of FORMULAS κ' that is consistent but not ω-consistent. κ' is consistent, since otherwise 17 Gen r would be κ-PROVABLE. However, κ' is not ω-consistent, because, by $\overline{\text{Bew}_\kappa}$(17 Gen r) and (15), $(x)\text{Bew}_\kappa Sb(r^{17}_{Z(x)})$ and, a fortiori, $(x)\text{Bew}_{\kappa'} Sb(r^{17}_{Z(x)})$ hold, while on the other hand, of course, $\text{Bew}_{\kappa'}$[Neg(17 Gen r)] holds.[46]

We have a special case of Theorem VI when the class κ consists of a finite number of FORMULAS (and, if we so desire, of those resulting from them by TYPE ELEVATION).

[45a] Since all existential statements occurring in the proof are based upon Theorem V, which, as is easily seen, is unobjectionable from the intuitionistic point of view.

[46] Of course, the existence of classes κ that are consistent but not ω-consistent is thus proved only on the assumption that there exists some consistent κ (that is, that P is consistent).

Every finite class κ is, of course, recursive.[46a] Let a be the greatest number contained in κ. Then we have for κ

$$x \,\varepsilon\, \kappa \sim (Em, n)[m \leqq x \,\&\, n \leqq a \,\&\, n \,\varepsilon\, \kappa \,\&\, x = m \; Th \; n].$$

Hence κ is recursive. This allows us to conclude, for example, that, even with the help of the axiom of choice (for all types) or the generalized continuum hypothesis, not all propositions are decidable, provided these hypotheses are ω-consistent.

In the proof of Theorem VI no properties of the system P were used besides the following:

1. The class of axioms and the rules of inference (that is, the relation "immediate consequence") are recursively definable (as soon as we replace the primitive signs in some way by natural numbers);
2. Every recursive relation is definable (in the sense of Theorem V) in the system P.

Therefore, in every formal system that satisfies the assumptions 1 and 2 and is ω-consistent there are undecidable propositions of the form $(x)F(x)$, where F is a [191] recursively defined property of natural numbers, and likewise in every extension of such a system by a recursively definable ω-consistent class of axioms. As can easily be verified, included among the systems satisfying the assumptions 1 and 2 are the Zermelo-Fraenkel and the von Neumann axiom systems of set theory,[47] as well as the axiom system of number theory consisting of the Peano axioms, recursive definition (by schema (2)), and the rules of logic.[48] Assumption 1 is satisfied by any system that has the usual rules of inference and whose axioms (like those of P) result from a finite number of schemata by substitution.[48a]

3

We shall now deduce some consequences from Theorem VI, and to this end we give the following definition:

A relation (class) is said to be *arithmetic* if it can be defined in terms of the notions $+$ and $.$ (addition and multiplication for natural numbers)[49] and the logical constants $\vee$, $\overline{}$, (x), and $=$, where (x) and $=$ apply to natural numbers only.[50] The notion "arithmetic proposition" is defined accordingly. The relations "greater than" and "congruent modulo n", for example, are arithmetic because we have

$$x > y \sim \overline{(Ez)}[y = x + z],$$
$$x \equiv y \pmod{n} \sim (Ez)[x = y + z.n \vee y = x + z.n].$$

[46a] [[On page 190, lines 21, 22, and 23, of the German text the three occurrences of α are misprints and should be replaced by occurrences of κ.]]

[47] The proof of assumption 1 turns out to be even simpler here than for the system P, since there is just one kind of primitive variables (or two in von Neumann's system).

[48] See Problem III in *Hilbert 1928a*.

[48a] As will be shown in Part II of this paper, the true reason for the incompleteness inherent in all formal systems of mathematics is that the formation of ever higher types can be continued into the transfinite (see *Hilbert 1925*, p. 184 [[above, p. 387]]), while in any formal system at most denumerably many of them are available. For it can be shown that the undecidable propositions constructed here become decidable whenever appropriate higher types are added (for example, the type ω to the system P). An analogous situation prevails for the axiom system of set theory.

[49] Here and in what follows, zero is always included among the natural numbers.

[50] The definiens of such a notion, therefore, must consist exclusively of the signs listed, variables for natural numbers, $x, y, \ldots$, and the signs 0 and 1 (variables for functions and sets are not permitted to occur). Instead of x any other number variable, of course, may occur in the prefixes.

We now have

Theorem VII. *Every recursive relation is arithmetic.*

We shall prove the following version of this theorem: every relation of the form $x_0 = \varphi(x_1, \ldots, x_n)$, where φ is recursive, is arithmetic, and we shall use induction on the degree of φ. Let φ be of degree s ($s > 1$). Then we have either [192]

1. $\varphi(x_1, \ldots, x_n) = \rho[\chi_1(x_1, \ldots, x_n), \chi_2(x_1, \ldots, x_n), \ldots, \chi_m(x_1, \ldots, x_n)]$[51]

(where ρ and all χ_1 are of degrees less than s) or

2. $\varphi(0, x_2, \ldots, x_n) = \psi(x_2, \ldots, x_n)$,

$$\varphi(k + 1, x_2, \ldots, x_n) = \mu[k, \varphi(k, x_2, \ldots, x_n), x_2, \ldots, x_n]$$

(where ψ and μ are of degrees less than s).

In the first case we have

$$x_0 = \varphi(x_1, \ldots, x_n) \sim (Ey_1, \ldots, y_m)[R(x_0, y_1, \ldots, y_m) \mathbin{\&} S_1(y_1, x_1, \ldots, x_n) \mathbin{\&} \ldots \mathbin{\&} S_m(y_m, x_1, \ldots, x_n)],$$

where R and S_i are the arithmetic relations, existing by the induction hypothesis, that are equivalent to $x_0 = \rho(y_1, \ldots, y_m)$ and $y = \chi_i(x_1, \ldots, x_n)$, respectively. Hence in this case $x_0 = \varphi(x_1, \ldots, x_n)$ is arithmetic.

In the second case we use the following method. We can express the relation $x_0 = \varphi(x_1, \ldots, x_n)$ with the help of the notion "sequence of numbers" (f)[52] in the following way:

$$x_0 = \varphi(x_1, \ldots, x_n) \sim (Ef)\{f_0 = \psi(x_2, \ldots, x_n) \mathbin{\&} (k)[k < x_1 \rightarrow f_{k+1} = \mu(k, f_k, x_2, \ldots, x_n)] \mathbin{\&} x_0 = f_{x_1}\}.$$

If $S(y, x_2, \ldots, x_n)$ and $T(z, x_1, \ldots, x_{n+1})$ are the arithmetic relations, existing by the induction hypothesis, that are equivalent to $y = \psi(x_2, \ldots, x_n)$ and $z = \mu(x_1, \ldots, x_{n+1})$, respectively, then

$$x_0 = \varphi(x_1, \ldots, x_n) \sim (Ef)\{S(f_0, x_2, \ldots, x_n) \mathbin{\&} (k)[k < x_1 \rightarrow T(f_{k+1}, k, f_k, x_2, \ldots, x_n)] \mathbin{\&} x_0 = f_{x_1}\}. \quad (17)$$

We now replace the notion "sequence of numbers" by "pair of numbers", assigning to the number pair n, d the number sequence $f^{(n,d)}$ $(f_k^{(n,d)} = [n]_{1+(k+1)d})$, where $[n]_p$ denotes the least nonnegative remainder of n modulo p.

We then have

Lemma 1. If f is any sequence of natural numbers and k any natural number, there exists a pair of natural numbers, n, d such that $f^{(n,d)}$ and f agree in the first k terms.

Proof. Let l be the maximum of the numbers $k, f_0, f_1, \ldots, f_{k-1}$. Let us determine an n such that

$$n \equiv f_i[\mathrm{mod}(1 + (i + 1)l!)] \quad \text{for } i = 0, 1, \ldots, k - 1,$$

which is possible, since any two of the numbers $1 + (i + 1)l!$ $(i = 0, 1, \ldots, k - 1)$ [193]

[51] Of course, not all $x_1, \ldots, x_n$ need occur in the χ_1 (see the example in footnote 27).

[52] f here is a variable with the [infinite] sequences of natural numbers as its domain of values. f_k denotes the $(k + 1)$th term of a sequence f (f_0 denoting the first).

62

are relatively prime. For a prime number contained in two of these numbers would also be contained in the difference $(i_1 - i_2)l!$ and therefore, since $|i_1 - i_2| < l$, in $l!$; but this is impossible. The number pair $n, l!$ then has the desired property.

Since the relation $x = [n]_p$ is defined by

$$x \equiv n \pmod{p} \;\&\; x < p$$

and is therefore arithmetic, the relation $P(x_0, x_1, \ldots, x_n)$, defined as follows:

$$P(x_0, \ldots, x_n) \equiv (En, d)\{S([n]_{d+1}, x_2, \ldots, x_n) \;\&\; (k)\,[k < x_1 \rightarrow T([n]_{1+d(k+2)}, k, [n]_{1+d(k+1)}, x_2, \ldots, x_n)] \;\&\; x_0 = [n]_{1+d(x_1+1)}\},$$

is also arithmetic. But by (17) and Lemma 1 it is equivalent to $x_0 = \varphi(x_1, \ldots, x_n)$ (the sequence f enters in (17) only through its first $x_1 + 1$ terms). Theorem VII is thus proved.

By Theorem VII, for every problem of the form $(x)F(x)$ (with recursive F) there is an equivalent arithmetic problem. Moreover, since the entire proof of Theorem VII (for every particular F) can be formalized in the system P, this equivalence is provable in P. Hence we have

Theorem VIII. *In any of the formal systems mentioned in Theorem VI*[53] *there are undecidable arithmetic propositions.*

By the remark on page 610, the same holds for the axiom system of set theory and its extensions by ω-consistent recursive classes of axioms.

Finally, we derive the following result:

Theorem IX. *In any of the formal systems mentioned in Theorem VI*[53] *there are undecidable problems of the restricted functional calculus*[54] (that is, formulas of the restricted functional calculus for which neither validity nor the existence of a
[194] counterexample is provable).[55]

This is a consequence of

Theorem X. *Every problem of the form $(x)F(x)$ (with recursive F) can be reduced to the question whether a certain formula of the restricted functional calculus is satisfiable* (that is, for every recursive F we can find a formula of the restricted functional calculus that is satisfiable if and only if $(x)F(x)$ is true.

By formulas of the restricted functional calculus (r. f. c.) we understand expressions formed from the primitive signs $\overline{}$, $\vee$, (x), $=$, $x, y, \ldots$ (individual variables), $F(x), G(x, y), H(x, y, z), \ldots$ (predicate and relation variables), where (x) and $=$ apply to individuals only.[56] To these signs we add a third kind of variables, $\varphi(x)$, $\psi(x, y)$,

[53] These are the ω-consistent systems that result from P when recursively definable classes of axioms are added.

[54] See *Hilbert and Ackermann 1928.*

In the system P we must understand by formulas of the restricted functional calculus those that result from the formulas of the restricted functional calculus of *PM* when relations are replaced by classes of higher types as indicated on page 599.

[55] In *1930a* I showed that every formula of the restricted functional calculus either can be proved to be valid or has a counterexample. However, by Theorem IX the existence of this counterexample is *not* always provable (in the formal systems we have been considering).

[56] Hilbert and Ackermann (*1928*) do not include the sign $=$ in the restricted functional calculus. But for every formula in which the sign $=$ occurs there exists a formula that does not contain this sign and is satisfiable if and only if the original formula is (see *Gödel 1930a*).

63

$\kappa(x, y, z)$, and so on, which stand for object-functions [Gegenstandsfunktionen] (that is, $\varphi(x)$, $\psi(x, y)$, and so on denote single-valued functions whose arguments and values are individuals).[57] A formula that contains variables of the third kind in addition to the signs of the r. f. c. first mentioned will be called a formula in the extended sense (i. e. s.).[58] The notions "satisfiable" and "valid" carry over immediately to formulas i. e. s., and we have the theorem that, for any formula A i. e. s., we can find a formula B of the r. f. c. proper such that A is satisfiable if and only if B is. We obtain B from A by replacing the variables of the third kind, $\varphi(x)$, $\psi(x, y)$, ..., that occur in A with expressions of the form $(\imath z)F(z, x)$, $(\imath z)G(z, x, y)$, ..., by eliminating the "descriptive" functions by the method used in *PM* (I, *14), and by logically multiplying[59] the formula thus obtained by an expression stating about each $F, G, \ldots$ put in place of some $\varphi, \psi, \ldots$ that it holds for a unique value of the first argument [for any choice of values for the other arguments].

We now show that, for every problem of the form $(x)F(x)$ (with recursive F), there is an equivalent problem concerning the satisfiability of a formula i. e. s., so that, on account of the remark just made, Theorem X follows.

Since F is recursive, there is a recursive function $\Phi(x)$ such that $F(x) \sim [\Phi(x) = 0]$, and for Φ there is sequence of functions, $\Phi_1, \Phi_2, \ldots, \Phi_n$, such that $\Phi_n = \Phi$, $\Phi_1(x) = x + 1$, and for every Φ_k $(1 < k \leq n)$ we have either

1. $$(x_2, \ldots, x_m)[\Phi_k(0, x_2, \ldots, x_m) = \Phi_p(x_2, \ldots, x_m)],$$
$$(x, x_2, \ldots, x_m)\{\Phi_k[\Phi_1(x), x_2, \ldots, x_m] = \Phi_q[x, \Phi_k(x, x_2, \ldots, x_m), x_2, \ldots, x_m]\}, \quad (18)$$
with $p, q < k$,[59a]

[195]

or

2. $$(x_1, \ldots, x_m)[\Phi_k(x_1, \ldots, x_m) = \Phi_r(\Phi_{i_1}(\mathfrak{x}_1), \ldots, \Phi_{i_s}(\mathfrak{x}_s))],\text{[60]} \quad (19)$$
with $r < k$, $i_v < k$ (for $v = 1, 2, \ldots, s$),

or

3. $$(x_1, \ldots, x_m)[\Phi_k(x_1, \ldots, x_m) = \Phi_1(\Phi_1, \ldots, \Phi_1(0))]. \quad (20)$$

We then form the propositions

$$(x)\overline{\Phi_1(x) = 0} \;\&\; (x, y)[\Phi_1(x) = \Phi_1(y) \rightarrow x = y], \quad (21)$$

$$(x)[\Phi_n(x) = 0]. \quad (22)$$

In all of the formulas (18), (19), (20) (for $k = 2, 3, \ldots, n$) and in (21) and (22) we now replace the functions Φ_i by function variables φ_i and the number 0 by an

[57] Moreover, the domain of definition is always supposed to be the *entire* domain of individuals.

[58] Variables of the third kind may occur at all argument places occupied by individual variables, for example, $y = \varphi(x)$, $F(x, \varphi(y))$, $G(\psi(x, \varphi(y)), x)$, and the like.

[59] That is, by forming the conjunction.

[59a] [The last clause of footnote 27 was not taken into account in the formulas (18). But an explicit formulation of the cases with fewer variables on the right side is actually necessary here for the formal correctness of the proof, unless the identity function, $I(x) = x$, is added to the initial functions.]

[60] The $\mathfrak{x}_i$ $(i = 1, \ldots, s)$ stand for finite sequences of the variables $x_1, x_2, \ldots, x_m$; for example, x_1, x_3, x_2.

individual variable x_0 not used so far, and we form the conjunction C of all the formulas thus obtained.

The formula $(Ex_0)C$ then has the required property, that is,

1. If $(x)[\Phi(x) = 0]$ holds, $(Ex_0)C$ is satisfiable. For the functions $\Phi_1, \Phi_2, \ldots, \Phi_n$ obviously yield a true proposition when substituted for $\varphi_1, \varphi_2, \ldots, \varphi_n$ in $(Ex_0)C$);

2. If $(Ex_0)C$ is satisfiable, $(x)[\Phi(x) = 0]$ holds.

Proof. Let $\psi_1, \psi_2, \ldots, \psi_n$ be the functions (which exist by assumption) that yield a true proposition when substituted for $\varphi_1, \varphi_2, \ldots, \varphi_n$ in $(Ex_0)C$. Let $\mathfrak{J}$ be their domain of individuals. Since $(Ex_0)C$ holds for the functions ψ_i, there is an individual a (in $\mathfrak{J}$) such that all of the formulas (18)–(22) go over into true propositions, (18′)–(22′), when the Φ_i are replaced by the ψ_i and 0 by a. We now form the smallest subclass of $\mathfrak{J}$ that contains a and is closed under the operation $\psi_1(x)$. This subclass ($\mathfrak{J}'$) has the property that every function ψ_i, when applied to elements of $\mathfrak{J}'$, again yields elements of $\mathfrak{J}'$. For this holds of ψ_1 by the definition of $\mathfrak{J}'$, and by (18′), (19′), and (20′) it carries over from ψ_i with smaller subscripts to ψ_i with larger ones. The functions that result from the ψ_i when these are restricted to the domain $\mathfrak{J}'$ of individuals will be denoted by ψ_i'. All of the formulas (18)–(22) hold for these functions also (when we replace 0 by a and Φ_i by ψ_i').

Because (21) holds for ψ_1' and a, we can map the individuals of $\mathfrak{J}'$ one-to-one onto the natural numbers in such a manner that a goes over into 0 and the function ψ_1' into the successor function Φ_1. But by this mapping the functions ψ_i' go over into the functions Φ_i, and, since (22) holds for ψ_n' and a, $(x)[\Phi_n(x) = 0]$, that is, $(x)[\Phi(x) = 0]$,
[196] holds, which was to be proved.[61]

Since (for each particular F) the argument leading to Theorem X can be carried out in the system P, it follows that any proposition of the form $(x)F(x)$ (with recursive F) can in P be proved equivalent to the proposition that states about the corresponding formula of the r. f. c. that it is satisfiable. Hence the undecidability of one implies that of the other, which proves Theorem IX.[62]

4

The results of Section 2 have a surprising consequence concerning a consistency proof for the system P (and its extensions), which can be stated as follows:

Theorem XI. *Let κ be any recursive consistent*[63] *class of* FORMULAS; *then the* SENTENTIAL FORMULA *stating that κ is consistent is not κ*-PROVABLE; in particular, the consistency of P is not provable in P,[64] provided P is consistent (in the opposite case, of course, every proposition is provable [in P]).

The proof (briefly outlined) is as follows. Let κ be some recursive class of FORMULAS chosen once and for all for the following discussion (in the simplest case it is the

[61] Theorem X implies, for example, that Fermat's problem and Goldbach's problem could be solved if the decision problem for the r. f. c. were solved.

[62] Theorem IX, of course, also holds for the axiom system of set theory and for its extensions by recursively definable ω-consistent classes of axioms, since there are undecidable propositions of the form $(x)F(x)$ (with recursive F) in these systems too.

[63] "κ is consistent" (abbreviated by "Wid(κ)") is defined thus: $\mathrm{Wid}(\kappa) \equiv (Ex)(\mathrm{Form}(x)\ \&\ \overline{\mathrm{Bew}_\kappa(x)})$.

[64] This follows if we substitute the empty class of FORMULAS for κ.

empty class). As appears from 1, page 608, only the consistency of κ was used in proving that 17 Gen r is not κ-PROVABLE;[65] that is, we have

$$\mathrm{Wid}(\kappa) \rightarrow \overline{\mathrm{Bew}_\kappa}(17 \text{ Gen } r), \tag{23}$$

that is, by (6.1),

$$\mathrm{Wid}(\kappa) \rightarrow (x)\, \overline{x\, B_\kappa\, (17 \text{ Gen } r)}.$$

By (13), we have

$$17 \text{ Gen } r = Sb(p^{19}_{Z(p)}),$$

hence

$$\mathrm{Wid}(\kappa) \rightarrow (x)\, \overline{x\, B_\kappa\, Sb(p^{19}_{Z(p)})},$$

that is, by (8.1), [197]

$$\mathrm{Wid}(\kappa) \rightarrow (x) Q(x), p). \tag{24}$$

We now observe the following: all notions defined (or statements proved) in Section 2,[66] and in Section 4 up to this point, are also expressible (or provable) in P. For throughout we have used only the methods of definition and proof that are customary in classical mathematics, as they are formalized in the system P. In particular, κ (like every recursive class) is definable in P. Let w be the SENTENTIAL FORMULA by which $\mathrm{Wid}(\kappa)$ is expressed in P. According to (8.1), (9), and (10), the relation $Q(x, y)$ is expressed by the RELATION SIGN q, hence $Q(x, p)$ by r (since, by (12), $r = Sb(q^{19}_{Z(p)})$), and the proposition $(x)Q(x\, p)$ by 17 Gen r.

Therefore, by (24), w Imp (17 Gen r) is provable in P[67] (and a fortiori κ-PROVABLE). If now w were κ-PROVABLE, then 17 Gen r would also be κ-PROVABLE, and from this it would follow, by (23), that κ is not consistent.

Let us observe that this proof, too, is constructive; that is, it allows us to actually derive a contradiction from κ, once a PROOF of w from κ is given. The entire proof of Theorem XI carries over word for word to the axiom system of set theory, M, and to that of classical mathematics,[68] A, and here, too, it yields the result: There is no consistency proof for M, or for A, that could be formalized in M, or A, respectively, provided M, or A, is consistent. I wish to note expressly that Theorem XI (and the corresponding results for M and A) do not contradict Hilbert's formalistic viewpoint. For this viewpoint presupposes only the existence of a consistency proof in which nothing but finitary means of proof is used, and it is conceivable that there exist finitary proofs that *cannot* be expressed in the formalism of P (or of M or A).

Since, for any consistent class κ, w is not κ-PROVABLE, there always are propositions (namely w) that are undecidable (on the basis of κ) as soon as $\mathrm{Neg}(w)$ is not κ-PROVABLE; in other words, we can, in Theorem VI, replace the assumption of ω-consistency by the following: The proposition "κ is inconsistent" is not κ-PROVABLE. (Note that there are consistent κ for which this proposition is κ-PROVABLE.) [198]

[65] Of course, r (like p) depends on κ.

[66] From the definition of "recursive" on page 602 to the proof of Theorem VI inclusive.

[67] That the truth of w Imp (17 Gen r) can be inferred from (23) is simply due to the fact that the undecidable proposition 17 Gen r asserts its own unprovability, as was noted at the very beginning.

[68] See *von Neumann 1927*.

In the present paper we have on the whole restricted ourselves to the system P, and we have only indicated the applications to other systems. The results will be stated and proved in full generality in a sequel to be published soon.[68a] In that paper, also, the proof of Theorem XI, only sketched here, will be given in detail.

Note added 28 August 1963. In consequence of later advances, in particular of the fact that due to A. M. Turing's work[69] a precise and unquestionably adequate definition of the general notion of formal system[70] can now be given, a completely general version of Theorems VI and XI is now possible. That is, it can be proved rigorously that in *every* consistent formal system that contains a certain amount of finitary number theory there exist undecidable arithmetic propositions and that, moreover, the consistency of any such system cannot be proved in the system.

[68a] [[This explains the "I" in the title of the paper. The author's intention was to publish this sequel in the next volume of the *Monatshefte*. The prompt acceptance of his results was one of the reasons that made him change his plan.]]

[69] See *Turing 1937*, p. 249.

[70] In my opinion the term "formal system" or "formalism" should never be used for anything but this notion. In a lecture at Princeton (mentioned in *Princeton University 1946*, p. 11 [[see *Davis 1965*, pp. 84–88]]) I suggested certain transfinite generalizations of formalisms, but these are something radically different from formal systems in the proper sense of the term, whose characteristic property is that reasoning in them, in principle, can be completely replaced by mechanical devices.

ON COMPLETENESS AND CONSISTENCY

(*1931a*)

Let Z be the formal system that we obtain by supplementing the Peano axioms with the schema of definition by recursion (on one variable) and the logical rules of the *restricted* functional calculus. Hence Z is to contain no variables other than variables for individuals (that is, natural numbers), and the principle of mathematical induction must therefore be formulated as a rule of inference. Then the following hold:

1. Given any formal system S in which there are finitely many axioms and in which the sole principles of inference are the rule of substitution and the rule of implication, if S contains[1] Z, S is incomplete, that is, there are in S propositions (in

[1] That a formal system S contains another formal system T means that every proposition expressible (provable) in T is expressible (provable) also in S.

[[Remark by the author, 18 May 1966:]]

[This definition is not precise, and, if made precise in the straightforward manner, it does not yield a sufficient condition for the nondemonstrability in S of the consistency of S. A sufficient condition is obtained if one uses the following definition: "S contains T if and only if every meaningful formula (or axiom or rule (of inference, of definition, or of construction of axioms)) of T *is a* meaningful formula (or axiom, and so forth) of S, that is, if S is an extension of T".

Under the weaker hypothesis that Z is recursively one-to-one translatable into S, with demonstrability preserved in this direction, the consistency, even of very strong systems S, *may* be provable in S and even in primitive recursive number theory. However, what can be shown to be unprovable in S is the fact that the rules of the equational calculus applied to equations, between primitive recursive terms, demonstrable in S yield only correct numerical equations (provided that S possesses the property that is asserted to be unprovable). Note that it is necessary to prove this "outer" consistency of S (which for the usual systems is trivially equivalent with consistency) in order to "justify", in the sense of Hilbert's program, the transfinite axioms of a

Notes

Chaper 1: Introduction

1. Ladriere, Limitations internes des formalismes, which is out of print. It includes a summary of mathematical results which point to limitations in deductive systems, and a chapter concerned with the problem of "capturing" mathematical intuition in a formal system.

2. Nagel and Newman, Gödel's Proof. This work includes some introductory material on the mathematical problems to which Godel addressed himself.

Chapter 2: An exposition of Gödel's Theorem

1. This is Gödel's original paper. Two English translations are available. One is by Meltzer and is published alone in a volume. Another is included in van Heijenoort's anthology, From Frege to Godel.

2. The definition appears in the original on pages 179 and 180. What Godel calls a recursive function is called a primitive recursive function today.

3. A number theoretic function is one from the natural numbers to the natural numbers.

4. Hatcher, in Foundations of Mathematics, requires a projection function as primitive, in addition to these two. Whether Gödel implicitly requires this function I do not know.

5. This is the Fundamental Theorem of Arithmetic. Proofs are available in many places. For one, see Herstein , I.N., Topics in Algebra, Blaisdell Publishing Co., Waltham, Mass., 1964, p.19.

6. Nagel and Newman use a different numbering technique based on a slightly different system.

7. See van Heijenoort for a reprint of Peano's paper.

8. Implication, equivalence, etc., are defined in the usual manner based on the primitive notions.

9. $\mathrm{Subst}\left(a\,{v \atop c}\right|$ means substituting c for the variable v in a.

10. A relation is recursive if there exists a recursive function satisfied by the members of the relation.

11. Gothic letters represent n-tuples of variables.

12. Other developments of Gödel's Theorem do not require ω-consistency. See Rosser, 1939.

13. Tarski has exhibited a system which has been shown to be simply consistent, but not ω-consistent. See Rosser, 1939.

Chapter 3: On Richard and Gödel

1. The paper is reprinted in van Heijenoort.

2. p. 63.

3. The version of Richard's Paradox given by Nagel and Newman is, of course, not the one originally given. R's work concerns set membership, whereas N&N involve predication of properties to individual numbers. The two are possibly reconcilable.

Chapter 4: An Extension to Ordinary Language

1. The paradox is developed in an unpublished paper of John Post, at Vanderbilt. The contradiction comes when acceptance of the original proposition entails a sentence which is necessarily true, and therefore not either possibly false, or neither true nor false.

Bibliography

The following bibliography is a condensed list of typical works on Godel in particular, but also includes works in mathematical logic and on the various paradox problems in the literature.

Alston, William P., The Philosophy of Language, Prentice-Hall, Englewood Cliffs, N.J., 1964

Baizin, Marcel, "Logique symbolique, Sur la portee du theoreme de M. Godel", in Acad. Roy. Belgique. Bull. el Sci, vol 26 (1940): 230-239.

Benacerraf, Paul and Hilary, Putnam, eds., Philosophy of mathematics: selected readings, Prentice-Hall, Englewood cliffs, N.J., 1964

Bernays, Paul, "Sur le platonisme dans les mathematiques", in L'ensignement mathematique, vol 34 (1935): 52-69

----------, "Quelque points essentiels de la metamathematique", in L'enseignement mathematique, vol. 34 (1935): 70-95.

Beth, Evert Willem, Les fondements logiques des mathematiques, Gauthier-Villars, Paris, 1955

----------, L'existence en mathematiques, Gauthier-Villars, Paris, Nauwelaerts, Louvain, 1956

----------, The Foundations of Mathematics, North Holland, Amsterdam, 1959. This is an English translation of the 1955 French work.

Carnap, Rudolf, Logical Syntax of Language, Humanities Press, New York, 1937

----------, Meaning and Necessity, University of Chicago Press, Chicago, 1947

Cavailles, Jean, Methode axiomatique et formalisme, essai sur le probleme du fondement des mathematiques, Hermann & C, Paris, 1938

Davis, Martin, ed., The Undecidable, Basic Papers on Undecidable Proposition, Unsolvable Problems, and Computable Functions, Raven Press, Hewlett, New York, 1965

Findlay, J., "Goedelian sentences: a non-numerical approach", in Mind, vol. 51 (1942):259-265.

Fraenkel, Abraham A., "Recent controversies about the foundation of mathematics", in Scripta Math., vol. 13 (1947):17-36

Geach, P.T., "Assertion", in Philosophical Review, vol. LXXIV(1965 no. 4); 449

Godel, Kurt, "Uber formal unentscheidbara Satze der Principia Mathematica und verwandter Systeme I", in Monatshefte fur Mathematik und Physik, vol. 38 (1931):173-198

---------, On formally undecidable propositions of Principia Mathematica and related systems, tr. B. Meltzer, intr. by R. B. Braithwaite, Oliver and Boyd, Ebinburgh, 1962

Hatcher, William S., Foundations of Mathematics, W. B. Saunders Co., Philadelphia, 1968

Heijenoort, J. van, From Frege to Godel, Harvard Univ. Cambridge, Mass., 1967

Kattsoff, Louis O., A Philosophy of Mathematics, Iowa College Press, Ames, Iowa, 1948

Kleene, S. C., Introduction to Metamathematics, New York, 1952

Ladriere, Jean, Les Limitations Internes des Formalisme, Paris,1957

Martin, Robert M., "Toward a Solution to the Liar Paradox", in Philosophical Review, LXXVI (1967 no. 3): 279

---------, "On Grelling's Paradox", in Philosophical Review, vol. LXXVII (1968 no. 3): 321

Mostowski, Andrzej, Sentences undecidable in formalized arithmetic: An exposition of the theory of Kurt Godel, North Holland, Amsterdam, 1952

Nagel, Ernest, and Newman, James R., Godel's Proof, New York University Press, New York, 1958

Post, John F., "The Modal Liar: Paradox in Self-Referential Falsifiability", Vanderbilt, unpublished paper

Quine, W.V.O, "On decidability and completeness", in Synthese, vol. 7(1949):441-446

----------, From a Logical Point of View, Harvard University Press, Cambridge, Mass., 1953

Richards, Thomas J., "Self-Referential Paradoxes", in Mind, vol. LXXVI (July 1967):303

Rosser, J.B., "Extensions of some theorems of Godel and Church", J. Sym. Logic, vol 1(1936):87-91

----------, "An informal exposition of proofs of Godel's Theorems and Church's Theorem", J. Sym. Logic, vol 4(1939):53-60

Russell, B. and Whitehead, A.N., Principia Mathematica, Cambridge Univ. Press, Cambridge, 1910

Tarski, A., "On undecidable statements in enlarged systems of logic and the concept of truth", in J. Sym. Logic, vol. 4(1939):105-112

----------, Logic, semantics, metamathematics, Oxford University Press, Oxford, 1956

Turing, A.M., "Computing machinery and intelligence", in Mind, vol 59(1950):433-460

Weyl, Hermann, Philosophy of Mathematics and Natural Science, tr. by Olaf Helmer, Princeton University Press, Princeton, 1949

Wilder, R.L., "The nature of mathematical proof", in Amer. Math. Monthly, vol 51(1944):309-323